ORSAY
n° d'ordre : 2715

UNIVERSITE PARIS-SUD

Centre D'Orsay

THESE

De Doctorat D'Etat Es Sciences Mathematiques

présentée pour obtenir le grade de

DOCTEUR ES-SCIENCES

par

Jean DESHAYES en collaboration avec Dominique PICARD

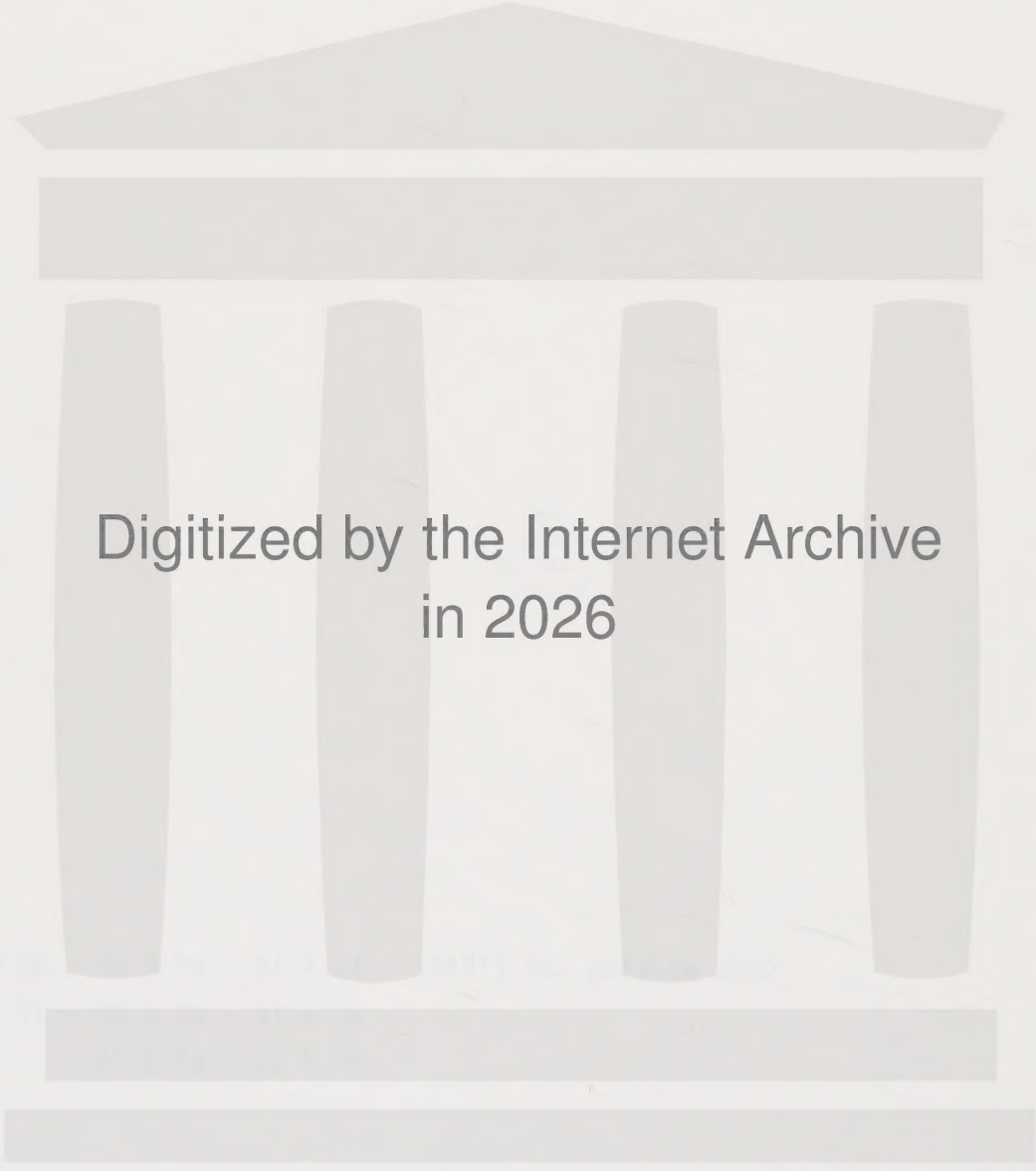
Sujet de la Thèse : *RUPTURES DE MODELES EN STATISTIQUE*

Soutenue le *10 mai 1983* ***devant le Jury composé de*** :

MM. D. DACUNHA-CASTELLE *Président*
R. AZENCOTT
A. BENVENISTE
J. BRETAGNOLLE
P. DEHEUVELS
C. GUEGUEN
J. SJÖSTRAND

UNIV
LONDON
MUS.

Code matière AMS (1980) : 62 A 10 - 62 E 20 - 62 F 05
62 F 12 - 62 F 20 - 62 G 10
62 J 05 - 62 M 15



Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41543524_0003

ORSAY
n° d'ordre :

UNIVERSITE PARIS-SUD

Centre D'Orsay

THESE

De Doctorat D'Etat Es Sciences Mathematiques

présentée pour obtenir le grade de

DOCTEUR ES-SCIENCES

par

Jean DESHAYES en collaboration avec Dominique PICARD

Sujet de la Thèse : *RUPTURES DE MODELES EN STATISTIQUE*

Soutenue le *10 mai 1983* ***devant le Jury composé de*** :

MM. *D. DACUNHA-CASTELLE* *Président*
R. AZENCOTT
A. BENVENISTE
J. BRETAGNOLLE
P. DEHEUVELS
C. GUEGUEN
J. SJÖSTRAND

The aim of this thesis is to present the careful asymptotic investigation of various methods to detect and estimate change-points in (time-ordered) statistical models. For miscellaneous statistical frameworks (parametric or non parametric models, independent or dependent observations, discrete or continuous time), it is developed a global introduction of classical methods, the description of new proceedings, just as a systematic comparison of these ones by means of simulations and an asymptotic investigation. The widely true optimality of likelihood methods led us to study them thoroughly : a theorem on the convergence of the double indexed likelihood process (change-point and model parameters) to a process depending on a brownian motion with a drift, provides a new and unified study : it particularly emphasizes the connexion between the amplitude and the instant parameters of the change-point, the limiting distribution of normalized test statistics and that of the maximum likelihood estimates.

En présentant ce travail, je tiens à exprimer ma reconnaissance à toutes les personnes qui m'ont permis de le mener à bien. Mes remerciements s'adressent d'abord à Didier Dacunha-Castelle : il m'a fait découvrir les probabilités-statistiques et continue depuis des années à approfondir mes connaissances ainsi qu'à guider mes travaux de recherche ; aux moments décisifs, ses encouragements ont été déterminants et ses propositions ont contribué à l'amélioration des pages qui suivent. Au fil des années, il a réussi à développer à Orsay une équipe et un environnement agréable de travail ; ma collaboration avec Guy Morel et Valentine Genon-Catalot a été bénéfique, même si elle n'est pas concrétisée ici.

Avec Dominique Picard, je ne peux faire le compte de ce que je lui dois ; son optimisme et sa persuasion ont soutenu notre travail ensemble durant ces cinq à six années, ses idées et son esprit de synthèse ont beaucoup contribué à l'élaboration des pages qui suivent : que ces mots lui renouvellent le témoignage de mon amitié.

Je remercie Patrice Assouad pour sa grande disponibilité : sa vaste culture et son écoute bienveillante nous ont été précieuses.

Mes remerciements vont également à Robert Azencott : les discussions avec lui (trop rares ?) sont toujours sources d'éclaircissement et sa rigueur source d'approfondissement.

Je veux exprimer ma reconnaissance à Jean Bretagnolle : ses exigences sur les plans mathématique et humain ainsi que sa technique exemplaire nous ont beaucoup apporté.

Je remercie aussi Albert Benveniste et Paul Deheuvels de manifester, par leur présence dans le jury, l'intérêt qu'ils portent à ce travail.

Johannes Sjöstrand a bien voulu à travers les fronts d'ondes analytiques m'entraîner sur le terrain des équations aux dérivées partielles : je tiens à le remercier pour sa gentillesse.

Pour m'avoir accueilli chaleureusement à l'E.N.S.T., je remercie très vivement Claude Gueguen et avec lui, tous les membres du département Systèmes et Communications : ils m'ont laissé une disponibilité suffisante pour mener cette thèse à son terme et m'offrent un environnement propice à la poursuite de ce travail.

Enfin, permettez-moi d'insister sur la dactylographie remarquable d'Anne-Marie Baillet et Nicole Parvan ainsi que sur leur amabilité car nous ne leur facilitons pas toujours la tâche ! Merci également à Mesdames Launay et Zielinski pour le soin qu'elles ont apporté à la réalisation finale.

INTRODUCTION

Lorsqu'on observe un phénomène durant une longue période, le problème de la stabilité du modèle est fréquemment posé. Lorsqu'on étudie la dépendance entre plusieurs variables, une seule régression linéaire ne suffit souvent pas pour une bonne modélisation. Cela conduit de façon naturelle à deux démarches différentes : envisager une famille de modèles plus large (par exemple, régression polynomiale, non-paramétrique...) ou bien conserver le modèle initial mais seulement sur une partie de l'espace d'observation (par exemple, régression linéaire par morceaux) ; dans ce dernier cas, se pose le problème de déterminer les points où se produisent les changements de modèle. D'autre part, des circonstances extérieures (crise économique, changement brusque des conditions climatiques, application d'un traitement, ...) amènent parfois une modification brutale des paramètres du modèle. La littérature statistique foisonne de problèmes où il y a rupture de modèles : citons-en quelques uns empruntés à différents domaines d'application.

L'expérimentation sur le blé [30] met en évidence une liaison linéaire entre le développement mesuré par le logarithme du nombre de primordia et le nombre de jours depuis le semis avec un instant de rupture situé au voisinage de l'épiaison.

La synthèse enzymatique pendant le cycle cellulaire varie continûment en fonction de la concentration de l'enzyme dans des cultures synchrones de cellules de levure de bière [60] mais pour les enzymes synthétisés au cours du cycle

cellulaire, la vitesse de synthèse double brutalement en certains points caractéristiques du cycle, restant constante avant et après ce point.

Les chercheurs sur la leucémie [36] qui étudient le temps de rechute après le début de rémission de la maladie sont amenés à tester l'hypothèse d'un taux de survie constant contre l'existence d'un point de rupture inconnu avec taux de survie constant avant et après.

L'étude de la hauteur de la couche stagnante dans un débit contrôlé de l'eau en aval d'un canal incliné ([1], [59]) conduit à poser une variation linéaire par morceaux du logarithme de la hauteur de la couche en fonction du logarithme du débit.

Dans un contrôle de qualité, un dérèglement des machines peut provoquer un changement brusque des caractéristiques de la production ou plus simplement augmenter la proportion d'objets défectueux par lot [39], il est alors important de détecter le plus vite possible s'il y a eu changement.

Notre travail est consacré à l'étude statistique de telles situations où des changements brusques de modèles se produisent, incluant les cas où les ruptures présentent certains degrés de continuité. Plusieurs points de vue sont envisageables et envisagés : modèles paramétriques ou non paramétriques, méthodes bayésiennes ou non bayésiennes, techniques séquentielles ou globales... . Pour mieux situer notre étude, présentons d'abord un bref résumé (pas forcément chronologique) de la littérature sur ce sujet.

Même si des exemples de régression linéaire avec changement de pente apparaissent dans la littérature avant 1950 (avec GARNER et HAMMOND en 1938, REEVE et HUXLEY en 1945 : voir [58]), c'est à partir de ces années-là que les problèmes statistiques commencent à être posés : avec RAO en 1948 [48] et surtout avec PAGE en 1954 [39] qui peut être considéré comme l'initiateur des ruptures de modèles avec des problèmes de contrôle de production : dans [39] il s'agit

de détecter un changement de moyenne dans une suite de variables indépendantes ; même si le critère statistique adopté n'était pas celui de WALD, la nécessité de déceler un changement éventuel de moyenne à chaque pas était évidente et la statistique de test proposée était la somme partielle des observations à un instant donné à laquelle on retranchait la plus petite des sommes partielles précédentes ([39], [40], [41]).

Pour un problème un peu différent : régression linéaire à deux phases avec instant de rupture inconnu, QUANDT ([45], [46]) propose d'utiliser le test du rapport de vraisemblance (variables supposées gaussiennes de variance éventuellement différente avant et après rupture). Après avoir conjecturé que la statistique de test égale à moins deux fois le logarithme du rapport de vraisemblance suivait asymptotiquement une loi de χ^2 (comme dans un modèle régulier) sous l'hypothèse nulle, il s'est aperçu par des simulations que la distribution dépendait très fortement du nombre d'observations et ne convergeait pas. Ceci tient à la particularité du paramètre inconnu : instant de rupture et pas seulement à sa nature discrète : nous y reviendrons plus loin.

Avec CHERNOFF et ZACKS (1964) [10] apparaissent les premières approches bayésiennes de l'estimation ce qui les amène à définir de nouvelles statistiques pour tester (de façon globale) une rupture de moyenne ; le comportement asymptotique de ces statistiques est développé par KANDER et ZACKS (1966) [29]. Des tests non paramétriques sont proposés pour des problèmes analogues par BHATTACHARYYA et JOHNSON (1968) [6].

Un pas décisif a été franchi avec les travaux d'HINKLEY qui établit la loi limite de l'instant de rupture pour un modèle de régression linéaire (en gaussiennes) à deux phases avec continuité ([21], [23]), mais aussi par des méthodes de fluctuations, il établit la loi limite sur $\mathbb{Z}$ de ce même instant de rupture pour un modèle avec changement de moyenne en loi gaussienne [24] et en loi de Bernoulli [26]. Les techniques employées sont basées sur les sommes partielles d'observations, ce qui permet une détection pas à pas d'une rupture éventuelle. Ce point de vue est adopté et mis en valeur pour la détection de rupture

dans un modèle de régression à l'aide des résidus récursifs : l'article de BROWN-DURBIN-EVANS (1975) constitue une bonne revue des différents tests utilisés dans ce domaine, sans en faire une comparaison théorique ni asymptotique.

Parallèlement, des problèmes de calculs effectifs d'estimateurs sont posés : particulièrement, la maximisation de la vraisemblance ou la méthode des moindres carrés sous contraintes donnent lieu aux techniques développées par de nombreux auteurs citons SPRENT [58], ROBISON [49], HUDSON [28].

Avec HINKLEY, la théorie asymptotique de l'estimation a fait une avancée alors que pour les tests, la loi asymptotique des statistiques de tests n'était pas toujours établie, il conjecturait que pour un modèle de régression avec continuité, la loi de la statistique de test du rapport de vraisemblance devait être voisine d'une loi de Fisher à 3 et $n-4$ degrés de liberté, si n désigne le nombre d'observations, et donc asymptotiquement un χ^2 à 3 degrés de liberté.

FEDER (1975) [14] propose une étude asymptotique des estimateurs des paramètres d'une régression par morceaux en supposant qu'il y a toujours continuité et aboutit à la normalité asymptotique de l'estimateur des instants de rupture, la vitesse de convergence dépendant de la régularité de la rupture (ordre de la première dérivée présentant un saut). Dans [15], il montre que la statistique du rapport de vraisemblance suit asymptotiquement une loi de χ^2 lorsque le modèle est "identifié" sous l'hypothèse nulle : "identifié" implique que le nombre de phases de la régression est le même sous l'hypothèse nulle et sous l'alternative et exclut en particulier le test d'absence de rupture contre l'existence de deux phases.

L'idée de considérer l'instant de rupture dans une suite d'observations comme un paramètre continu est apparue et développée les années suivantes : BAGSHAW-JOHNSON [2], BHATTACHARYYA-BROCKWELL [7] testent des changements de moyenne

sur le mouvement brownien. HAWKINS [19] établit des relations de récurrence pour évaluer la loi de la statistique du rapport de vraisemblance dans un modèle gaussien avec rupture de moyenne ; il établit la non convergence asymptotique de la statistique à cause de la loi du logarithme itéré pour le mouvement brownien. L'étude de l'instant de rupture comme paramètre continu aboutit à l'approximation des processus de sommes partielles par un mouvement brownien et à des théorèmes d'invariance utilisés par Mc NEIL ([34], [35]) et SEN ([52], [53]).

Cependant, plusieurs questions cruciales restaient en suspens : la loi asymptotique de la statistique de vraisemblance n'était toujours pas établie, aussi bien pour des modèles avec rupture de moyenne, que pour des modèles de régression par morceaux avec ou sans continuité ; peu de comparaisons des différents tests avait été faites si ce n'est des simulations (citons GARBADE [17], et aussi BASSEVILLE [4] pour une revue). La loi asymptotique des estimateurs n'était établie que pour les modèles de régression avec continuité et pour les ruptures de moyenne en variables de Bernoulli ou gaussiennes ; de plus dans ces deux derniers cas, la loi limite de l'estimateur est d'emploi difficile car elle nécessite des calculs de probabilités différents pour chaque valeur de l'amplitude de la rupture.

Nous ne nous sommes pas intéressés au point de vue bayésien développé par exemple par SMITH [57] ni aux préoccupations récursives ou algorithmiques qui sous-tendent les travaux de SCHWEDER [51], PRABHU [44], WILLSKY [61], BASSEVILLE-BENVENISTE [4], ces derniers étant consacrés plus spécifiquement à la détection de rupture dans les systèmes.

Nous avons commencé notre étude des modèles avec rupture par une comparaison des tests proposés pour déceler un changement de paramètre dans une régression, et avons poursuivi par une étude approfondie des méthodes de vraisemblance : lois asymptotiques des statistiques de test et des estimateurs sur différents modèles. Notre travail est présenté en cinq parties.

Première Partie :

Bien que ce ne soit pas le début de notre étude, nous proposons pour des raisons pédagogiques en première partie le rapport (en anglais) qui a été l'objet d'une conférence invitée au Congrès Européen des Statisticiens de Brighton en 1980. Nous envisageons deux situations différentes où on cherche à détecter une rupture de modèle. La première situation est de type non paramétrique : pour tester un changement de loi dans une suite de n variables indépendantes, identiquement distribuées, indicées chronologiquement, nous étudions une version pas à pas du test de Kolmogorov-Smirnov comparant les fonctions de répartition empirique avant et après l'éventuelle rupture. Le bon cadre asymptotique consiste à rendre continu le paramètre : instant de rupture k en se ramenant à un paramètre $t = \frac{k}{n}$ compris entre 0 et 1. Le comportement asymptotique sous l'hypothèse nulle de ce test est bien compris à l'aide d'un théorème de convergence du processus empirique à double indice dont l'un est l'instant de rupture éventuel t. Celui-ci met en évidence la non convergence de la statistique du suprémum à cause des problèmes de rupture possible en début et en fin d'observation : t voisin de 0 ou 1. Le processus renormalisé, c'est-à-dire multiplié par $\sqrt{t(1-t)}$ converge alors vers un pont brownien à double indice, ce qui établit la loi limite de la statistique du suprémum avec "renormalisation".

La seconde situation envisagée est de type paramétrique : changement de moyenne ou de variance dans un modèle gaussien, plus généralement changement de paramètre dans une suite de n variables indépendantes, identiquement distribuées. Nous établissons un premier théorème d'invariance sur le logarithme du processus de vraisemblance qui consiste à considérer, comme dans le cadre non paramétrique, un processus à double indice dont l'un est le temps. Ce point de vue multi-indice fait bien comprendre les problèmes de bords et de bon paramétrage rencontrés antérieurement : les conséquences seront revues de façon détaillée dans les troisième et quatrième parties.

Dans cette première partie, figure aussi une comparaison asymptotique au sens non local des tests : le test du rapport de vraisemblance est optimal au sens de la vitesse de décroissance exponentielle des erreurs pour les modèles gaussiens et modèles exponentiels. Ce point est développé plus largement dans la seconde partie, mais cette étude est complètée par les résultats de simulations pour des tailles d'échantillons $n = 200$ et $n = 30$ qui confortent largement la comparaison asymptotique.

Deuxième partie :

Celle-ci est le point de départ de notre travail. Pour le modèle de régression linéaire sur variables indépendantes en loi gaussienne, nous comparons les différents tests proposés par BROWN-DURBIN-EVANS [9] qui sont basés sur les résidus récursifs : test des sommes cumulées des résidus récursifs ou de leurs carrés, test du rapport de vraisemblance. Le calcul effectif des vitesses de décroissance exponentielle des erreurs de première et second espèce utilise fondamentalement des théorèmes de grandes déviations pour marches aléatoires et aboutit à l'optimalité du test de vraisemblance : pour une vitesse exponentielle de décroissance du niveau fixée, la vitesse exponentielle de décroissance de l'erreur de 2^e espèce est strictement meilleure en tout point de la contre-hypothèse ; cela a comme conséquence d'attribuer au test du rapport de vraisemblance un meilleur "pouvoir séparant".

Si d désigne la différence entre les paramètres avant et après rupture, les vitesses de décroissance des erreurs de 2^e espace ainsi que les courbes définissant le "pouvoir séparant" s'expriment à l'aide de l'expression globale : $t(1-t)d^2$ qui semble être le véritable paramètre définissant une "taille" de rupture, quelle que soit sa position.

Nous avons aussi calculé les vitesses exponentielles pour un modèle où il est possible de faire un échantillonnage du temps. La comparaison de ce test avec le test du rapport de vraisemblance est fonction de la loi d'échantillonnage de départ et de la place de la rupture par rapport aux bords (t voisin de 0

ou 1). Cette seconde partie est présentée dans sa version condensée parue dans Teoriya Ver. Prim. (1982).

Troisième partie :

Ayant montré la supériorité du test du rapport de vraisemblance au sens non local, pour l'appliquer, il est nécessaire de connaître la loi de sa statistique sous l'hypothèse nulle ; en fait, nous sommes intéressés par sa loi asymptotique.

Comme cela est partiellement montré dans la première partie, c'est un point de vue processus à double-indice et théorème d'invariance qui permet d'éclairer les différents problèmes. Nous montrons ici un théorème de convergence du processus partiel de vraisemblance (double indice dont l'un est t) pour des variables indépendantes identiquement distribuées dans l'espace des fonctions continues à support compact. Les conditions suffisantes de régularité du modèle para métrique sont relativement larges et analogues à celles supposées habituellement pour la convergence du processus global de vraisemblance (mono-indice). La convergence au dehors des compacts de cette suite de processus nécessaire pour les tests conduit à la définition d'une nouvelle norme sur l'espace des fonctions continues à cause du problème de bord (t voisin de 0).

Ce théorème d'invariance est l'outil fondamental de notre étude et pourra être montré pour des modèles plus généraux que le cadre classique étudié ici.

Quatrième partie :

Elle concerne les applications aux tests d'existence d'une rupture et la loi des estimateurs des différents paramètres de la rupture (ou des ruptures). Pour un modèle statistique classique : n-échantillon et pour un modèle de régression linéaire par rapport au temps (ou toute autre variable observée à intervalles réguliers), le théorème d'invariance met en évidence les renormalisations possibles de la statistique de test du rapport de vrai-

semblance qui ne converge pas à cause des problèmes de bords : pour t voisin de 0 ou 1, le comportement fait intervenir la loi du logarithme itéré pour le processus limite qui s'exprime à l'aide du mouvement brownien. Nous proposons deux renormalisations : l'une dite interne, l'autre externe qui consiste dans les deux cas à pondérer les bords ; aux bords il est naturel de ne déceler que des ruptures de grande amplitude : réapparaît ce paramètre global $t(1-t)d^2$ quantifiant une rupture. Le théorème d'invariance permet en fait de règler tous les problèmes de test de vraisemblance : rupture de modèle avec ou sans continuité, rupture de moyenne seule, rupture de moyenne et de variance ... .

La seconde application du théorème d'invariance est l'étude globalisante des problèmes d'estimation. Pour contourner les difficultés de calculs de lois limites rencontrées par HINKLEY ([24], [26]), nous faisons une étude asymptotique dans le cas où l'amplitude de la rupture d_n tend lentement vers 0 ; cela permet d'aboutir à une loi limite universelle pour l'instant de rupture t qui est estimé à la vitesse nd_n^2 . Le théorème d'invariance permet également de traiter toutes les lois limites des estimateurs du maximum de vraisemblance pour des ruptures présentant différents degrés de continuité et aussi pour des ruptures intervenant près des bords.

Enfin, nous avons un éclairage sur la question non résolue antérieurement : que se passe-t-il pour les estimateurs lorsque le nombre de phases supposé est surestimé ? Le comportement sous l'hypothèse nulle du processus du rapport de vraisemblance implique que l'estimateur de l'instant de rupture tend vers 0 ou 1, que l'un des estimateurs des paramètres avant ou après rupture estime bien le vrai paramètre du modèle alors que l'autre tend vers l'infini.

Cinquième partie :

Nous envisageons des ruptures dans un cas particulier de modèle à temps continu et non stationnaire : le processus de Poisson non homogène. Le rapprochement entre le modèle du processus de Poisson avec rupture et le modèle de n-échantillon avec saut de la densité est intéressant. La détection d'une rupture pour un processus de Poisson homogène donne lieu à la comparaison de deux tests non paramétriques classiques. La loi asymptotique de l'estimateur de l'instant de rupture est établie pour une amplitude d fixe (cas d'Hinkley) et pour une amplitude tendant vers 0.

L'étude asymptotique des méthodes de vraisemblance analogue à celle de la quatrième partie est menée jusqu'au bout pour les processus de Poisson d'entensité paramétrique et nous envisageons l'extension de ces méthodes à des processus ponctuels plus généraux.

BIBLIOGRAPHIE

[1] BACON D.W. - WATTS D.G. (1971)
" Estimating the transition between two intersecting straight lines ". Biometrika, vol. 58, p. 525-534.

[2] BAGSHAW M. - JOHNSON R.A. (1975)
" Sequential detection of drift change in a Wiener process ". Comm. Stat. vol. 4, p. 787-796.

[3] BAGSHAW M. - JOHNSON R.A. (1977)
" Sequential procedures for detecting parameter changes in a time series model ". J.A.S.A. vol. 72, p. 593-597.

[4] BASSEVILLE M. - BENVENISTE A. (1981)
" Ruptures de modèles statistiques ". Rapport IRISA n° 145, Rennes.

[5] BECKMAN R.J. - COOK R.D. (1979)
" Testing for two-phase regressions ". Technometrics vol. 21, p. 65-69.

[6] BHATTACHARYYA P.K. - JOHNSON R.A. (1968)
" Nonparametric tests for shift at an unknown time point ". A.M.S. vol. 39, p. 1731-1743.

[7] BHATTACHARYYA P.K. - BROCKWELL P.J. (1976)
" The minimum of an additive process with applications to signal estimation and storage theory ". Z. Wahr. verw. Gebiete vol. 37, p. 51-75.

[8] BROWN R.L. - DURBIN J. (1968)
" Methods of investigating wether a regression relationship is constant over time ". Selected Statistical Papers. European Meeting. Mathematical Centre Tracts n° 26. Amsterdam.

[9] BROWN R.L. - DURBIN J. - EVANS J.M. (1975)
" Techniques for testing the constancy of regression relationships over time (with discussion) ". J.R.S.S. B vol. 37, p. 149-192.

[10] CHERNOFF H. - ZACKS S. (1964)
" Estimating the current mean of a normal distribution which is subjected to changes in time ". A.M.S. vol. 35, p. 999-1018.

[11] CHALMOND B. (1979)
" Ruptures de modèles pour les processus autorégressifs ". Thèse de 3e cycle. Orsay.

[12] COBB G.W. (1978)
" The problem of the Nile : conditional solution to a change-point problem ". Biometrika vol. 65, p. 243-251.

[13] FARLEY J.U. - HINICH M.J. (1970)
" A test for a shifting slope coefficient in a linear model ". J.A.S.A. vol. 65, p. 1320-1329.

[14] FEDER P.I. (1975)
" On asymptotic distribution theory in segmented regression problems-identified case ". Ann. Stat. vol. 3, p. 49-83.

[15] FEDER P.I. (1975)
" The log-likelihood ratio in segmented regression ". Ann. Stat. vol. 3, p. 84-97.

[16] FRIEDMAN M. (1982)
" Piecewise exponential models for survival data with covariates ". Ann. Stat. vol. 10, p. 101-113.

[17] GARBADE K. (1977)
" Two methods for examining the stability of regression coefficients ". J.A.S.A. vol. 72, p. 54-63.

[18] HAWKINS D.M. (1976)
" Point estimation of the parameters of piecewise regression models ". Applied Statistics vol. 25, p. 51-57.

[19] HAWKINS D.M. (1977)
" Testing a sequence of observations for a shift in location ". J.A.S.A. vol. 72, p. 180-186.

[20] HAWKINS D.M. (1980)
" A note on continuous and discontinuous segmented regression ". Technometrics vol. 22, p. 443.

[21] HINKLEY D.V. (1969)
" Inference about the intersection in two-phase regression ". Biometrika vol. 56, p. 495-504.

[22] HINKLEY D.V. (1970)
"Inference about the change-point in a sequence of random variables ". Biometrika vol. 57, p. 1-17.

[23] HINKLEY D.V. (1971)
" Inference in two-phase regression ". J.A.S.A. vol. 66, p. 736-743.

[24] HINKLEY D.V. (1971)
" Inference about the change-point from cumulative sum tests ". Biometrika vol. 58, p. 509-523.

[25] HINKLEY D.V. (1972)
" Time-ordered classification ". Biometrika vol. 59, p. 509-523.

[26] HINKLEY D.V. - HINKLEY E.A. (1970)
" Inference about the change-point in a sequence of binomial variables Biometrika vol. 57, p. 477-488.

[27] HINKLEY D.V. - CHAPMAN P. - RUNGER G. (1980)
" Change-point problems ". Techn. Report (University of Minnesota).

[28] HUDSON D.J. (1966)
" Fitting segmented curves whose join points have to be estimated ". J.A.S.A. vol. 61, p. 1097-1129.

[29] KANDER Z. - ZACKS S. (1966)
" Test procedures for possible changes in parameters of statistical distributions occuring at unknown time points ". A.M.S. vol. 37, p. 1196-1210.

[30] LERMAN P.M. (1980)
" Fitting segmented regression models by grid search ".
Appl. Stat. vol. 29, p. 77-84.

[31] Mc CABE B.P.M. - HARRISON M.J. (1980)
" Testing the constancy of regression relationships over time using least square residuals ". Appl. Stat. vol. 29, p. 142-148.

[32] Mc GEE V.E. - CARLETON W.T. (1970)
" Piecewise regression ". J.A.S.A. vol. 65, p. 1109-1124.

[33] Mc GILCHRIST C.A. - WOODYER K.D. (1975)
" Note on a distribution-free cusum techniques ".
Technometrics vol. 17, p. 321-325.

[34] Mc NEIL I.B. (1974)
" Tests for change of parameter at unknown time and distributions of some related functionals on Brownian motion ". Ann. Stat. vol. 2, p. 950-962.

[35] Mc NEIL I.B. (1978)
" Properties of sequences of partial sums of polynomial regression residuals with applications to tests for change of regression at unknown times ". Ann. Stat. vol. 6, p. 422-433.

[36] MATTHEWS D.E. - FAREWELL V.T. (1982)
" On testing for a constant hazard against a change-point alternative ".
Biometrics vol. 38, p. 463-468.

[37] OGATA Y. - AKAIKE H. - KATSURA K. (1982)
" The application of linear intensity models to the investigation of caudal relations between a point process and another stochastic process "
Ann. Inst. Stat. Math. vol. 34 B, p. 373-387.

[38] OURGHANLIAN B. (1981)
" Estimations et tests en rupture de régression ". Thèse de 3^e cycle.
Orsay.

[39] PAGE E.S. (1954)
" Continuous inspection schemes ". Biometrika vol. 41, p. 100-116.

[40] PAGE E.S. (1955)
" A test for a change in a parameter occuring at an unknown point ".
Biometrika vol. 42, p. 523-527.

[41] PAGE E.S. (1957)
" On problems in which a change of parameters occurs at an unknown point ". Biometrika vol. 44, p. 248-252.

[42] PETTITT A.N. (1979)
" A nonparametric approach to the change-point problem ".
Appl. Stat. vol. 28, p. 126-135.

[43] PETTITT A.N. (1980)
" A simple cumulative sum type statistic for the change-point problem with zero-one observations ". Biometrika vol. 67, p. 79-84.

[44] PRABHU K.P.S. (1970)
" On the detection of a sudden change in system parameters ".
I.E.E.E. Trans. on Inf. theory vol. 16, p. 497-500.

[45] QUANDT R.E. (1958)
" The estimation of the parameters of a linear regression system obeying two separate regimes ". J.A.S.A. vol. 53, p. 873-880.

[46] QUANDT R.E. (1960)
" Tests of the hypothesis that a linear system obeys two separate regimes ". J.A.S.A. vol. 55, p. 324-330.

[47] QUANDT R.E. (1972)
" A new approach to estimating switching regressions ".
J.A.S.A. vol. 67, p. 306-310.

[48] RAO C.R. (1948)
" The utilization of multiple measurements in problems of biological classification ". J.R.S.S. B n° 10, p. 159.

[49] ROBISON D.E. (1964)
" Estimates for the points of intersection of two polynomial regressions '
J.A.S.A. vol. 59, p. 214-224.

[50] ROYSTON J.P. - ABRAMS R.M. (1980)
" An objective method for detecting shift in basal body temperature in women ". Biometrics vol. 36, p. 217-224.

[51] SCHWEDER T. (1976)
" Some optimal methods to detect structural shift or outliers in regression ". J.A.S.A. vol. 71, p. 491-501.

[52] SEN P.K. (1980)
" Asymptotic theory of some tests for a possible change in the regression slope occuring at an unknown time point ". Z. Wahr. verw. Gebiete vol. 52, p. 203-218.

[53] SEN P.K. (1981)
" The Cox regression model, invariance principles for some induced quantile processes and some repeated significance tests ".
Ann. Stat. vol. 9, p. 109-121.

[54] SEN A.K. - SRIVASTAVA M.S. (1973)
" On multivariate tests for detecting change in mean ".
Sankhya A vol. 35, p. 173-186.

[55] SEN A.K. - SRIVASTAVA M.S. (1975)
" On tests for detecting change in mean ". Ann. Stat. vol. 3, p. 98-108.

[56] SEN A.K. - SRIVASTAVA M.S. (1975)
" Some one-sided tests for change in level ". Technometrics vol. 17, p. 61-64.

[57] SMITH A.F.M. (1975)
" A bayesian approach to inference about a change-point in a sequence of random variables ". Biometrika vol. 62, p. 407-416.

[58] SPRENT P. (1961)
" Some hypotheses concerning two phase regression lines ". Biometrics vol. 17, p. 634-645.

[59] TISHLER A. - ZANG I. (1981)
" A new maximum likelihood algorithm for piecewise regression ". J.A.S.A. vol. 76, p. 980-987.

[60] WILLIAMS D.A. (1970)
" Discrimination between regression models to determine the pattern of enzyme synthesis in synchronous cell cultures ". Biometrics vol. 26.

[61] WILLSKY A.S. (1976)
" A survey of several failure detection methods ". Automatica, p. 601-611.

[62] WORSLEY K.J. (1979)
" On the likelihood ratio test for a shift in location of normal parameters ". J.A.S.A. vol. 74, p. 365-367.

TESTING FOR A CHANGE-POINT IN STATISTICAL MODELS

J. DESHAYES, D. PICARD.(*)

SUMMARY : *To test for one change-point in a statistical model we have considered different usual techniques and introduced new ones in parametric and non parametric situations. To solve the tricky problem of asymptotic distributions (likelihood ratio test, Kolmogorov-Smirnov process, ...), we adopted the point of view of writing the test statistics as functionals of double indiced processes (the second indice being the unknown change-point) for which limit theorems are obtained. After looking at both local and non local asymptotic comparisons , the latter seemed to be more tractable : it gave interesting results such as optimality of some likelihood ratio tests and optimization of the boundaries for the cusum techniques. Moreover, its conclusions seemed to be largely accorded to simulation results.*

KEYWORDS : *Kolmogorov-Smirnov statistics, cusum techniques, likelihood techniques, weak convergence of double-indiced processes, asymptotic local and non local comparison of tests, large deviations.*

(*) *Université Paris-Sud, Département Mathématique*
E.R.A. 532 "Statistique Appliquée"
91405 ORSAY CEDEX (France)

0 - INTRODUCTION

We consider the problem of testing for a change point in a statistical model : suppose we observe n variables $X_1,\ldots,X_n$. Usually we assume to have a sample of a distribution function F belonging to a family $\mathcal{F}$. That means three assumptions:

I $X_1,\ldots,X_n$ are independant

II $X_1,\ldots,X_n$ are identically distributed

III The distribution function of each X_i belongs to $\mathcal{F}$.

Here we are just going to question about the second assumption. That is, we are going to test the hypothesis H_o : $X_1,\ldots,X_n$ are independant, identically distributed variables ; versus H_1 : there exists a change point k such that, the random variables $X_1,\ldots,X_k$ are i.i.d., but after k, the variables are still independant, still with a distribution function belonging to $\mathcal{F}$, but not identically :

H_o : $X_1,\ldots,X_n$ are i.i.d. with distribution $F_o \in \mathcal{F}$.

H_1 : $X_1,\ldots,X_n$ are independant

$$\exists k\ ,\ 1 < k < n \text{ such that } \begin{cases} X_1,\ldots,X_k \text{ are i.i.d. } \sim F \\ X_i \sim F_i \neq F\ ,\ i > k \end{cases}$$

and a particular very useful case is a change point with stationarity before and after k : it will be noticed by : H_1^* :

$$H_1^* : \exists k,\ 1 < k < n \text{ such that } \begin{cases} X_1,\ldots,X_k \text{ are i.i.d. } \sim F \\ X_{k+1},\ldots,X_n \text{ are i.i.d.} \sim F^* \end{cases}$$

$$F,\ F^* \in \mathcal{F} \quad .$$

Many practical examples have been developped in the litterature ([8], [12], [20], [19], [1], [17], [27], [26]...). Most of them fit in one of the two following models in which the initial problem can be reduced to the previous situation.

Example I : Change-point in a regression model: suppose we observe the random variables y_i linearly depending on the regressor vector x_i :

$$y_i = x_i' \beta_i + \epsilon_i \qquad i = 1,\ldots,n$$

x_i : r-dimensional column vector of the regressors
β_i : r-dimensional column vector of the parameters
ϵ_i : errors : i.i.d. $N(0, \sigma^2)$

and we want to test the constancy of the parameters β_i :

$$H_o : \{\beta_i = \beta_o \quad \forall i\} \text{ versus}$$

$$H_1 : \{\exists k, \ 1 < k < n, \quad \beta_i = \beta \ , \ i \leq k$$
$$\beta_i \neq \beta \ , \ i > k$$

and we'll particularly study the stationary alternative

$$H_1^* : \{\exists k, \ 1 < k < n, \quad \beta_i = \beta, \ i \leq k, \ \beta_i = \beta^*, \ i > k\}$$

With the introduction of the recursive residuals w_i, which were applied in the change point problem for the first time in Brown, Durbin and Evan's paper [8], the initial problem can be reduced to the previous situation where F_o is fixed (or almost fixed if σ^2 is unknown) and equals to $N(0, \sigma^2)$ and the family $\mathcal{F}$ is equal to the set of all normal distribution functions with variance σ^2 .

$$w_i = \frac{y_i - x_i' B_{i-1}}{\sqrt{1+x_i' (X_{i-1}' X_{i-1})^{-1} x_i}} \qquad i = r+1,\ldots,n$$

where $X_i' = (x_1,\ldots, x_i)$

$Y_i = (y_1,\ldots,y_i)$

$B_i = (X_i' X_i)^{-1} X_i' Y_i$

under H_o, the w_i are i.i.d. $N(0, \sigma^2)$

under H_1, the w_i are i.i.d. $N(0, \sigma^2)$ for $i \leq k$

independant, $N(\mu_i, \sigma^2)$ for $i > k$

(for instance, under H_1^* $\mu_i = \frac{x_i' (X_{i-1}' X_{i-1})^{-1} X_k' X_k (\beta^*-\beta)}{\sqrt{1 + x_i' (X_{i-1}' X_{i-1})^{-1} x_i}}$

Example II : Change point in a Poisson Process.

Suppose we observe a Poisson Process $N(t)$, with the intensity $\nu(t)$ for t belonging to the interval $[0,1]$, and we want to test the stationarity of ν against one change at a point τ with stationarity before and after :

$$H_o = \{\nu(t) = \nu_o \qquad \forall t \in [0,1]\}$$

$$H_1^* = \{\exists\, \tau \in]0,1[\text{ such that } \nu(t) = \nu \text{ for } t \leq \tau$$
$$\nu(t) = \nu^* \text{ for } t > \tau\}$$

If we notice n for $N(1)$ and k for $N(\tau)$, with the introduction of the observed times $X_1,\ldots,X_n$ between the consecutive jumps $T_1,\ldots,T_n$, we almost reduce the problem to the previous situation, with the family $\mathcal{F}$ being the set of exponential distribution functions, but with both the number of observations n and the change point k being random variables.

$n=N(1)$ $k=N(\tau)$ $T_1,\ldots,T_{n-1}$ waiting-times defined by

$T_i = \inf \{ t \in [0,1] \ / \ N(t) \geq i \}$.

$$X_1 = T_1, \ldots, X_i = T_i - T_{i-1} \ , \ X_n = T_n - T_{n-1}$$

In the first part of the paper, we'll give a survey of different types of tests, the second part will be devoted to asymptotic comparison between some of them (both local and non local points of view will be investigated). The third section is concerned with computations of tests powers for different values of n, by Monte Carlo methods and the last section is devoted to theoritical developments on multi-indice point of view.

1 - PICKING UP TESTS

The litterature about change point problems is too much extensive to envisage giving a complete list of all possible tests. We just want to examine here the simplest ways to construct change point tests starting from classical situations of comparison between two populations.

Indeed, in a change point problem if you suppose the exact place k of the change is known, then you are in a classical situation of comparison between two populations and the methods you are going to employ depend on the information you dispose about the data essentially through two points :

- the first point consists on the knowledge of F_o (as is the example I where it is known that F_o is equal to $N(0,\sigma^2)$) : the question is then a problem of goodness of fit for the n-k last observations to the law F_o (the k first observations being then harmless), or the unknowledge of F_o (and then the problem becomes a real comparison between two populations).

- the second point is the size of the family $\mathcal{F}$, that leads to parametric or non parametric procedures. When the change point is unknown there are two classical ways of extending usual tests to that situation.

The first one is to take as statistic of test an average of the statistics obtained with fixed k, with respect to an arbitrary wheighting on the nuisance parameter k .

This point of view more or less of Bayesian type has been developped in many papers, for instance [3], [11], [28].

In this paper we shall be mostly concerned with the second way which consists in taking as critical region the union of all the critical regions obtained for the different values of k . But, very often, such procedures, appearing as functionals of a partial sums process, offer a great difficulty for approximating their level, which we propose to circumvent by adopting a multi-indice point of view (adding the unknown time-change indice k).

The last section is entirely concerned with this multi-indic point of view, begining with a general invariance theorem, which allows a generalization of a large class of statistical limit theorems to multi-indice ones (such as convergence of the Kolmogorov statistic to a brownian bridge, or likelihood ratio process to a degenerated gaussian one).

For each test, we give, besides the construction, first usual properties (consistancy, elementary remarks...). First let's look at the non parametric procedures :

I.1. NON PARAMETRIC TESTS.

a) Kolmogorov-Smirnov : Suppose that $\mathcal{F}$ is a set of continuous c.d.f. defined on $\mathbb{R}$. When F_o is known, the classical test with critical region :

$$\sup_x \sqrt{n} \left| \hat{F}_n(x) - F_o(x) \right| > c$$

(where $\hat{F}_n(x)$ is the empirical cumulative distribution function), can be used with either known or unknown change point k .

But if k is known, a better test can be used, calculating the e.c.d.f. $\hat{F}^*_{n-k}(x)$ only on the n-k last observations :

$$\sup_x \sqrt{n-k} \left| \hat{F}^*_{n-k}(x) - F_o(x) \right| > c$$

the extension to the case k unknown leads to critical region :

$$\sup_k \sup_x \sqrt{n} \left(1 - \frac{k}{n}\right) \left| F^*_{n-k}(x) - F_o(x) \right| > c$$

These two tests have the advantage of being more responsive and especially more suitable to be extended to the case F_o unknown.

For that case, we propose for critical region :

$$\sup_k \sup_x \sqrt{n} \left(1 - \frac{k}{n}\right) \frac{k}{n} \left| \hat{F}^*_{n-k}(x) - \hat{F}_k(x) \right| > c$$

where $\hat{F}_k(x)$ is the e.c.d.f. calculated on the k first observations and takes the place of $F_o(.)$. Notice that, in each case, the statistic is distribution-free under null hypothesis and normalizing factors have been introduced ($\frac{k}{n}$, $\frac{k}{n}\left(1 - \frac{k}{n}\right)$) to get convergence of the associated processes.

To approximate the level of these tests,

one possibility is to get the convergence of the concerned processes in $\mathscr{C}([0,1]^2)$, with uniform topology. Let us define the process $Y_n(t,u)$, on $[0,1]^2$ as follows :

- For all fixed k in $\{1,2,\ldots,n\}$

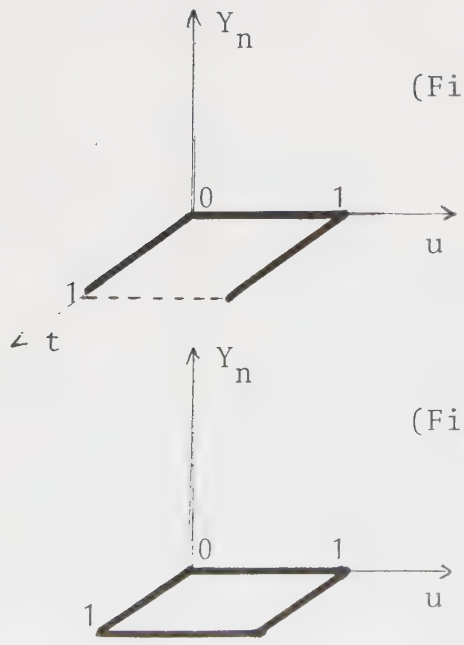
$$Y_n\left(\frac{k}{n}, u\right) = \sqrt{n}\,\frac{k}{n}\left(G_n\left(\frac{k}{n},u\right)-u\right)$$

where $G_n\left(\frac{k}{n},.\right)$ is the polygonal line constructed on the points $\{(0,0), (X^k_{(i)}, \frac{i}{k+1}), (1,0)\}$ where $X^k_{(i)}$ is the order statistic in the k-sample $X_{n-k+1},\ldots,X_n$.

- For all fixed $u \in [0,1]$.

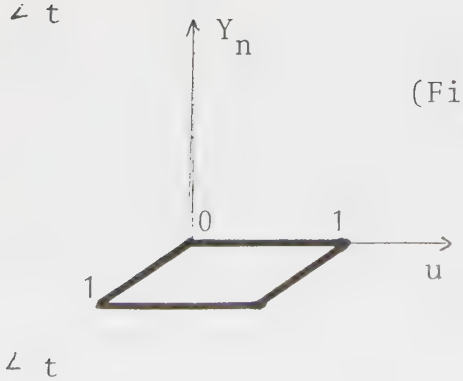
$Y_n(.,u)$ is polygonal line constructed on the points $\{(0,0), (\frac{k}{n}, Y_n(\frac{k}{n},u)), k=1,\ldots,n, (1,0)\}$.

Using the results of Komlos Major Tusnady [32], we can deduce the asymptotic levels of the first two tests from the convergence of the empirical process $\{Y_n(t,u), (t,u) \in [0,1]^2\}$ to the double indiced Wiener Process, tied down on the 3 underlined sides : $\{Y(t,u), (t,u) \in [0,1]^2\}$ (Fig. I)



(Fig. I)

and the last level is obtained from convergence of $Y'_n(t,u) = Y_n(t,u) - t\,Y_n(1,u)$ to the double indiced Brownian bridge (Fig. II)



(Fig. II)

The consistancy of these tests obviously depends on the alternatives and is not always true : we give two examples where the problem of consistancy can be solved .

Example (i) : Suppose that k is in fact a function $k(n)$ and the cumulative distribution functions $F_{i,n}$, after the change are such that :

$$\frac{1}{n-k(n)} \sum_{i=k(n)+1}^{n} F_{i,n}(x) \xrightarrow[n\to\infty]{} G(x) \qquad \forall x \in \mathbb{R}$$

(Take for instance a translation model with a convergent sequence of parameters :

$$F_{i,n}(X) = F(x-\theta_{i,n}) \qquad \theta_{i,n} = 0 \qquad i \leqslant k(n)$$

$$\theta_{i,n} \xrightarrow[n\to\infty]{} \theta \qquad i > k(n)$$

Then if $G \neq F$ and $\overline{\lim}_{n\to\infty} \frac{k(n)}{n} < 1$, the test is consistant.

If $G = F$, non consistancy can be deduced from results of Dvoretzky, Kiefer, Wolfowitz [16], Bretagnolle [6] or Le Cam (unpublished).

Example (ii) : When the change point happens to be at the edges of the sample, it is rather intuitive that one of the two populations to be compared, is badly estimated, and the test can be unconsistant :

If $\frac{k(n)}{\sqrt{n}}$ (or $\frac{n-k(n)}{\sqrt{n}}$ in the case F_o unknown) tends to 0 when n tends to infinity, the tests are not consistant.

b) Mann-Whitney's test :

An identical extension of the classical Mann Whitney's test to the change point problem has been developped by Pettitt [25], [26] using the statistic :

$$\sup_k \sum_{i=1}^{k} \sum_{j=k+1}^{n} \operatorname{sgn}(X_i - X_j)$$

As in classical situation, comparison between M.W. and K.S. procedures seemed to strongly depend on the family $\mathcal{F}$.

I.2. PARAMETRIC METHODS : CUSUM TECHNIQUES.

Now, we are going to investigate more parametric methods and first, moments methods. As usual, let us start from the case k known and F_o known (Possibly only through its two first moments).

A possibility of testing the constancy of the model is to compare the theoritical mean to the empirical one calculated on the n-k last observations. Assume (for simplification) that :

$$E_{F_o} X = 0 \qquad E_{F_o} X^2 = 1 \quad :$$

Put : $z^*_{n-k} = \sum_{i=k+1}^{n} \frac{X_i}{n-k}$

the critical region is defined by :

$$\{ |z^*_{n-k}| > c \}$$

To extend the test to the case k unknown (F_o still known) a possibility is to compare each $|z^*_j|$ to a function $g_n(\frac{j}{n})$ and reject H_o if one pass through the frontiere :

$$\mathcal{D} = \{\exists s \in [0,1] \ / |Z_n(s)| > g(s)\}$$

where : $Z_n(.)$ is the polygonal line constructed on the points $\{(0,0), (\frac{j}{n}, \frac{j}{\sqrt{n}} z_j^*) \quad j=1,\ldots,n\}$

g is a non negative continuous function defined on $[0,1]$.

The process $\{Z_n(s), s \in [0,1]\}$ is normalized to obtain the convergence to a Wiener Process $\{W(t), t \in [0,1]\}$ under H_o when n tends to infinity.

We have essentially two imperatives :

1) How to evaluate the level (here we'll essentially think of "approximate the level").

2) If possible to optimize the boundary.

The convergence of Z_n to the Wiener Process gives an answer for the functions f such that

(*) $$P(\exists t \in]0,1[\ / \ |W(t)| > g(t)) \neq 0 \text{ and } 1.$$

Only few functions g lead to explicit level and practically, tables exist for straight lines and that is probably why only these are used in practice . Yet it is not hard to obtain simulation-results for general functions verifying (*).

For the second question, answers will be given with the asymptotic non local comparison : it will be shown that a parabola would be the optimal boundary. Unfortunately it doesn't verify (*).

Although, as it will be the case for the likelihood ratio test, we'll be obliged to compromize and ask to the context or the practicians'needs to choose a boundary.

But first, let us look to some simple properties about power and consistancy leading to obvious improvements.

We shall restrict to the particular case of a translation model: $F_i(x) = F(x-\theta_i)$. The power is then determined by the function : $m_n(s)$, $s \in [0,1]$ where

$$m_n(s) = E_{H_1} Z_n(s)$$

and a sufficient condition for consistancy is :

$$\exists s \in]0,1[\ / \ |m_n(s)| \xrightarrow[n\to\infty]{} + \infty$$

Let us now have a look on typical shapes of the function $m_n(s)$.

i) In a translation model with stationarity after the change point :

$$m_n(s) \simeq \begin{cases} \theta^* \, s \sqrt{n} & \text{if} \quad s \leq 1 - \frac{k(n)}{n} \\ \theta^* \, (1 - \frac{k(n)}{n}) \sqrt{n} & \text{if} \quad s > 1 - \frac{k(n)}{n} \end{cases}$$

under $H_1^* = \{X_1, \ldots, X_{k(n)} \sim F(., 0)$

$$X_{k(n)+1}, \ldots, X_n \sim F(., \theta^*)\}$$

It can be noticed that if $k(n)$ behave as $n\tau$, $\tau \in]0,1[$, the rate of convergence of $m_n(s)$ is $\sqrt{n}$, and if $\frac{n-k(n)}{\sqrt{n}}$ tends to 0 when n tends to infinity the test is always unconsistan as it was for Kolmogorov Smirnov test.

ii) If we now turn to the regression model and suppose that the regressors x_i :

either equal to 1

or behave as i.i.d. random variables (and independant from the errors ε_i)

then $$m_n(s) \simeq \begin{cases} \Delta\tau \, \log \frac{1}{1-s} \sqrt{n}, & s \leq 1-\tau \\ \Delta\tau \, \log \frac{1}{\tau} \sqrt{n}, & s > 1-\tau \end{cases}$$

under $$H_1^*(\tau,\Delta) = \begin{cases} y_i = x_i' \beta + \varepsilon_i & i \le k(n) \\ y = x_i' \beta^* + \varepsilon_i & i > k(n) \end{cases}$$

$$\frac{k(n)}{n} \xrightarrow[n\to\infty]{} \tau \qquad \tau \in]0,1[$$

and $$\Delta = (Ex_i')\,(\beta^* - \beta)$$

In this model, this CUSUM test is nearly one of those proposed by Brown, Durbin, Evans [8]. In fact they used the process in reversed order. Because of the shape of $m_n(s)$, ($m_n(s) \ge m_n(1) - m_n(1-s)$, $\forall s$), the latter test is less powerful but has the advantage of being recursive. Actually this fact was yet pointed out in [8] and has been later developped in Schweder [27] .

With the same procedure as in non-parametric techniques, we can generalize to the case F_o unknown. Properties are quite the same, besides the fact that the process :

$Z_n'(.)$, defined on the points

$$\{(0,0),\ (\frac{k}{n}, \sqrt{n}\,\frac{k}{n}\,(1-\frac{k}{n})\,[\frac{\sum_{1}^{k} X_i}{k} - \frac{\sum_{k+1}^{n} X_i}{n-k}])\ k=1,\ldots,n\}$$

converges to a Brownian Bridge $\{B_o(t),\ t \in [0,1]\}$ instead of a Wiener Process.

Same constructions can also be done with higher moments.

I.3. LIKELIHOOD METHODS

We are now going to deal with parametric situations. $\mathcal{F}$ is supposed to be dominated by a measure μ . Let $f(.,\theta)$ denote the density (with respect to μ) of F_θ , θ belongs to Θ included in $\mathbb{R}$. One problem which has not get been solved is that of approximating the level of the Likelihood Ratio test. In this section, we are going to present a point of view, which is helpful for understanding theoritical edges-problems and affords a reasonnable solution of compromising for getting tests.

First we are going to summarize the results on convergence of the likelihood ratio process.

(Details and proofs are left in section 4, and especially explicite forms of conditions $A_1,\ldots, A_4$.

Theorem 1 :

Under the regularity conditions $A_1 \ldots A_4$ for Λ being a compact interval in $\mathbb{R}$ including 0, the double indiced Likelihood Process : $\{Z_{n,\theta_o}(t,\lambda),(t,\lambda)\in[0,1]\times\Lambda\}$ converges under $H_o(\theta_o)$ to the Gaussian θ-degenerated Process : $\{Z_{\theta_o}(t,\lambda),(t,\lambda)\in[0,1]\times\lambda\}$ where :

- For each λ, $Z_{n,\theta_o}(.,\lambda)$ is the polygonal line constructed on the points

$$\{(0,0),\ (\frac{k}{n}\ ,\ \sum_{i=1}^{k} \text{Log}\ \frac{f\ (X_i,\ \theta_o + \frac{\lambda}{\sqrt{n}})}{f\ (X_i,\ \theta_o)})\ k=1\ldots n\}$$

- $$Z_{\theta_o}(t,\lambda) = \lambda \sqrt{I(\theta_o)}\ W(t) - t\,\lambda^2\,\frac{I(\theta_o)}{2}$$

Corollary 1 :

Under the same assumptions $A_1 \ldots A_4$ and if for each k in $\{1,\ldots,n\}$ equation (A) : $\sum_{i=n-k+1}^{n} f'/f(X_i,\theta) = 0$ has an unique solution $\hat{\theta}_k^*$, then : under H_o, the one dimensional Process

$$Z^*_{n,\theta_o}(t) = 2t\, Z_{n,\theta_o}(t, \hat{\theta}^*_{[nt]}) \quad , \; t \in [0,1]$$

converges, when n goes to infinity, to the square of a Wiener Process :

$$\{W^2(t) \; , \; t \in [0,1]\} \quad .$$

Corollary 2 :

Under the same assumptions $A_1 \ldots A_4$ and if for each k in $\{1 \; ,\ldots,n\}$, the equations

$$A \; : \; \sum_{i=k+1}^{n} \frac{f'}{f}(X_i,\theta) = 0$$

$$B \; : \; \sum_{i=1}^{k} \frac{f'}{f}(X_i,\theta) = 0$$

have an unique solution $\hat{\theta}^*_{n-k}$ (resp $\hat{\theta}_k$) .

Then :

The one dimensional process :

$\{Z_n(t), \; t \in [\, 0,1 \,]\}$ converges, under H_o and when n goes to infinity to

$\{B^2_o(t) \; , \; t \in [\, 0,1 \,]\}$ where :

- $B_o(.)$ is the Brownian Bridge
- $Z_n(.)$ is the polygonal line constructed on the points :

$$\{(0,0), \; (\frac{k}{n}, \frac{2k}{n}(1-\frac{k}{n}) \left[\sum_{i=1}^{k} \text{Log} \frac{f(X_i,\hat{\theta}_k)}{f(X_i,\hat{\theta}_n)} + \sum_{i=k+1}^{n} \text{Log} \frac{f(X_i,\hat{\theta}^*_{n-k})}{f(X_i,\hat{\theta}_n)} \right]),$$

$$k = 1,\ldots,n \; .$$

The two statistics $\frac{n}{2k} Z^*_{n,\theta_o}(\frac{k}{n})$ and $\frac{n^2}{2k(n-k)} Z_n(\frac{k}{n})$ are

the likelihood ratio statistics for testing H_o against H_1^* respectively in the case θ_o known and θ_o unknown with a chang eventually appearing at the time n-k (resp. k). Going on with likelihood techniques would be taking the supremum over all possi ble change points k .

However the level-problem remains unsolved.

Thus we propose to circunmvent the edges problems by using a wheighting of the LR statistics especially unfavoura- ble for edges values of k , with the two tests :

(1) $$\mathscr{D}_3 = \{ \sup_{k \in \{1,\ldots,n\}} Z^*_{n,\theta_o}(\tfrac{k}{n}) > C_1 \} \qquad \theta_o \text{ known}$$

(2) $$\mathscr{D}'_3 = \{ \sup_{k \in \{1,\ldots,n\}} Z_n(\tfrac{k}{n}) > C_2 \} \qquad \theta_o \text{ unknown}$$

These two tests are obviously consistant at each point of H_1^* such that the change point k(n) satisfy :

$$0 < \underline{\lim}\ \frac{k(n)}{n} \quad (\text{resp } 0 < \underline{\lim}\ \frac{k(n)}{n} \leq \overline{\lim}\ \frac{k(n)}{n} < 1)$$

and optimality properties can be obtained as is going to be shown in the next section.

2 - Asymptotic comparison

1. LOCAL POINT OF VIEW.

The theory consists in comparing, for tests of $H_{o,n}$ against $H^*_{1,n}$ with same asymptotic level, the asymptotic type II errors under sequences of hypotheses $H_{1,n}(k_n, \theta_{1,n}, \theta_{2,n})$ which are contiguous to the sequence $H_{o,n}$. We shall prove that local asymptotic uniform optimality is impossible. To this end, we shall restrict to the elementary case of a simple hypothesis (θ_o known) against a family of contiguous alternatives of special type :

$$H^*_{1,n}(\tau, \theta_1) : \begin{cases} X_1, \ldots, X_{n\tau} \sim F_{\theta_o} \\ \\ X_{n\tau+1}, \ldots, X_n \sim F_{\theta_o + \frac{\theta_1}{\sqrt{n}}} \end{cases},$$

An interesting point in this lack of optimality is that it is not the fact of edges problems.

In that case, the "LR" process : $\{Z_{n,\theta_o}(t),\ t \in [0,1]\}$ is asymptotically sufficient and we may reasonnably restrict to "convergent" sequences of tests only depending on it, (Def. 1), in order to replace the research of an optimal test by an equivallent problem in continuous time.

Definition 1 : A sequence of test ϕ_n is said "convergent sequence only depending on $Z_{n,\theta_o}(.)$" if : There exist a sequence of measurable functionnals δ_n :

δ_n : $\mathscr{C}[0,1] \to [0,1]$ and a measurable function

δ : $\mathscr{C}[0,1] \to [0,1]$ such that :

i) $\delta_n \xrightarrow[n\to\infty]{} \delta$ uniformly on $\mathscr{C}[0,1]$

ii) $P(D_\delta) = 0$ where

D_δ is the set of discontinuities of δ

P is the Wiener measure.

iii) $\phi_n = \delta_n (Z_{n,\theta_o}(.))$

Lemma : Under the asumptions of corollary 2 (§ I) and restricting to "convergent sequences of tests only depending on $Z_{n,\theta_o}(.)$" we have the following proposition :

If there exists an optimal test for testing $H_o = \{X_1 \ldots X_n \sim F_{\theta_o}\}$ versus the family of alternatives :

$$H^*_{1n}(\tau,\theta_1) = \begin{cases} X_1 \ldots X_{n\tau} \sim F_{\theta_o} & , \quad \theta_1 \in \Theta_1 \\ X_{n\tau+1} \ldots X_n \sim F_{\theta + \frac{\theta_1}{\sqrt{n}}} & , \quad \tau \in [0,1\ ] \end{cases}$$

Then you can construct an optimal test, for the limit problem

Observing a Process $\{Z(t),\ t \in [0,1]\}$,

Test $H'_o = \{Z(t) = W(t)^2 ,\ t \in [0,1]\}$

versus :

$$H_1' = \{ (W(t) + g_{\tau,\theta_1}(t))^2 \quad , \quad t \in [\ ,1] \}$$

where $$g_{\tau,\theta_1}(t) = \theta_1 \ (1-\tau) \ 1_{t>1-\tau} + \theta_1 \ t \ 1_{t\leq 1-\tau}$$

Proposition : Under the conditions of the previous lemma. There is no optimal test for testing H_o versus $H_1^*(\tau,\theta_1)$ for every θ_1 belonging to a compact set of $\mathbb{R}^+$ and every τ belonging to $[0,1\ ]$.

Proof of the lemma : It is shown in Section 4 that the process :

$\{ Z_{n,\theta_o}(t) \quad , \quad t \in [0,1] \}$ converges - if θ_o is true

to $\{ W^2(t) \quad , \quad t \in [0,1] \}$

- if $H_1^*(\tau,\theta_1)$ is true

to $\{ (W(t)+g_{\tau,\theta_1}(t))^2 \ , \ t \in [0,1] \}$,

The properties (i) and (ii) of the sequences of tests set us in the conditions of theorem 5.1 of Billingsley, p. 30 [4] which concludes the proof.

Proof of the proposition : Cameron-Martin's formula gives the density of the measure associated with the process $Z_{\theta_o}(.)$ under H_1' ($Z_{\theta_o}(t) = W(t) + g_{\tau,\theta_1}(t), \ t \in [0,1]$) with respect to the Wiener measure :

$$\exp \{ \theta_1 \ Z_{\theta_o}(1-\tau) - \frac{\theta_1^2 \ (1-\tau)}{2} \}$$

It can be easily deduced that the likelihood ratio :

$$\frac{dP_{H_1'}}{dP_{H_o}} = \exp - \frac{\theta_1^2 \ (1-\tau)}{2} \ ch \{ \theta_1 \ Z^{1/2}(1-\tau) \}$$

On this expression, it becomes obvious that uniform optimality is impossible, since the family has no monotone likelihood ratio.

That point of view gives as information that an invariance reduction would be quasi-hopeless. However, one may be less exacting on uniform optimality and look for tests having an average asymptotic power uniformly optimal for θ (the average being taken over the space of the nuisance parameter k with respect to an arbitrary wheighting on the set $\{1,\ldots,n\}$) as has been done in a non parametric casing in Bhattacharya-Jonhson's paper [3] .

2. NON LOCAL POINT OF VIEW.

When the hypotheses, instead of being contiguous are get fixed, most of the presented tests can be adapted to have their both errors exponentially decreasing when n tends to infinity.

In that framework, several possibilities of comparisons can be developped.

The first one, which has been introduced by Chernoff [10], and then extensively developped in L.D. Brown [7], Birgé [5], Tusnady [22],..., has the essential interest of giving easily optimal procedures in various casings and it can be used to find optimal boundaries in moments methods and optimality properties for MLR Tests but is often considered as of poor practical interes in comparing two sequences of tests.

The second one, which can be regarded as very much related t Bahadur's slopes[2], will be used to more sensitive comparisons between some of the previous procedures.

a) Chernoff's comparison :

Among tests of $H_o : \{\theta \in \Theta_o\}$ against $H_1 : \{\theta \in \Theta_1\}$

$\Theta_0 \cap \Theta_1 = \emptyset$, we shall compare here the log-errors, when the sample size n goes to infinity.

Among tests $\{\phi_n\}$ having the same log-level α :

ie : $\sup_{\theta \in \Theta_0} \overline{\lim_{n\to\infty}} \frac{1}{n} \log E_{\theta_0} \phi_n \leq -\alpha$.

We compare the log-type II-errors and say

Definition :

The sequence $\{\phi_n\}$ is said : asymptotically more powerful than the sequence $\{\phi'_n\}$ at the log-level α, and at the point θ_1 of Θ_1 if :

$$\sup_{\theta\in\Theta_0} \overline{\lim_{n\to\infty}} \frac{1}{n} \log E_\theta \phi_n \leq -\alpha$$

$$\sup_{\theta\in\Theta_0} \overline{\lim_{n\to\infty}} \frac{1}{n} \log E_\theta \phi'_n \leq -\alpha$$

and $\overline{\lim_{n\to\infty}} \frac{1}{n} \log E_{\theta_1}(1-\phi_n) = -\overline{\gamma}(\theta_1) < -\underline{\gamma}'(\theta_1) = \underline{\lim_{n\to\infty}} \frac{1}{n} \log E_{\theta_1}(1-\phi'_n)$

This investigation techniques permit in particular cases of exponential families to find optimal tests.

We are going to explicit 4 examples where the optimal test is the LR Test (after reduction by invariance).

Example 1 : Change point of the mean in a normal family with known variance :

Proposition :

The LR Test is asymptotically uniformly most powerful among translation-invariant tests for testing

$H_0 = \{X_1 \ldots X_n$ i.i.d. $N(\theta, 1)\}$ versus

$$H_1^*(\tau,\Delta) = \{X_1 \ldots X_{k(n)} \text{ i.i.d. } N(\theta,1)$$

$$X_{k(n)+1},\ldots,X_n \text{ i.i.d. } N(\theta+\Delta,1)\}$$

where $\frac{k(n)}{n} \xrightarrow[n\to\infty]{} \tau \qquad \tau \in]0,1[$

$\Delta \in \mathbb{R} \setminus \{0\}$

Example 2 : Change point of the mean in a normal family with unknown variance.

Proposition : The LR Test is asymptotically uniformly most powerful among homothety-translation invariant tests for testing :

$$H_o = \{X_1,\ldots,X_n \text{ i.i.d. } N(\theta,\sigma^2)\} \text{ versus}$$

$$H_1^*(\tau,\Delta) = \{X_1 \ldots X_{k(n)} \text{ i.i.d. } N(\theta,\sigma^2)$$

$$X_{k(n)} \ldots X_n \text{ i.i.d. } N(\theta+\Delta,\sigma^2)$$

where $\frac{k(n)}{n} \xrightarrow[n\to\infty]{} \tau \qquad \tau \in]0,1[$

$\Delta \in \mathbb{R} \setminus \{0\}$

$\sigma^2 \in \mathbb{R}^+ \setminus \{0\}$

These two propositions can be extended without change to the case of a change point in a regression model if the regressors x_i are supposed to behave as i.i.d. random variables (cf. Deshayes-Picard [15]).

Example 3 : Change point of the variance in a normal family with known mean.

Proposition : The LR test is asymptotically uniformly most powerful among homothety-invariant tests for testing :

$$H_o = \{X_1 \ldots X_n \quad \text{i.i.d. } N(0,\sigma^2)\} \qquad \text{versus}$$

$$H_1^*(\tau,\lambda) = \{X_1,\ldots,X_{k(n)} \quad \text{i.i.d. } N(0,\sigma^2)$$

$$X_{k(n)+1},\ldots,X_n \quad \text{i.i.d. } N(0,\lambda\sigma^2)$$

where $\frac{k(n)}{n} \xrightarrow[n\to\infty]{} \tau \qquad \tau \in]0,1[\ , \ \lambda \in \mathbb{R}^+ \setminus \{0,1\}$

$$\sigma^2 \in \mathbb{R}^+ \setminus \{0\} \ .$$

Example 4 : Change point of the parameter in a model of exponential-law.

Proposition : The LR Test is asymptotically uniformly most powerful among homothety-invariant tests for testing :

$$H_o = \{X_1,\ldots,X_n \quad \text{i.i.d. with density } \theta\{\exp\theta x\}1_{X>0}\}$$

versus $H_1^*(\tau,\lambda) = \{X_1 \ldots X_{k(n)} \quad \text{i.i.d. } \theta\{\exp\theta\ x\}1_{X>0}$

$$X_{k(n)+1},\ldots,X_n \quad \text{i.i.d. } \lambda\theta\{\exp\lambda\theta x\}1_{X>0}\}$$

where $\frac{k(n)}{n} \xrightarrow[n\to\infty]{} \tau \qquad \tau \in]0,1[\ , \ \lambda \in \mathbb{R}^+ \setminus \{0,1\}$

$$\theta \in \mathbb{R}^+ \setminus \{0\}$$

Remark : This last case is a consequence of the Example 3 in which the sample sizes n and $k(n)$ are even.

Proof : The proof of all cases are identical, by noticing that, in the case k known, the reduction to maximal invariant : $T_{k,n}$ respectively :

i) $$\sqrt{\frac{k(n-k)}{n}}\left(\sum_{i=1}^{k}\frac{X_i}{k} - \sum_{i=k+1}^{n}\frac{X_i}{n-k}\right) \ ,$$

ii) $$\sqrt{\frac{k(n-k)}{n}}\ \frac{\left(\sum_{i=1}^{k}\frac{X_i}{k}-\sum_{i=k+1}^{n}\frac{X_i}{n-k}\right)}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\frac{1}{n}\sum_{i=1}^{n}X_i\right)^2}}$$

iii) $$\frac{\frac{1}{n}\sum_{i=1}^{k}X_i^2}{\frac{1}{n-k}\sum_{i=k+1}^{n}X_i^2}\ \Bigg)$$

leads to monotone likelihood ratio models. In each case, the critical regions

$$T_{n,k} < a_{n,k}\ ,\ T_{n,k} > b_{n,k}$$

provide optimal unilateral invariant test (for every fixed n). For α positive real, put $a_{n,k}(\alpha)$ and $b_{n,k}(\alpha)$ such that

$$\lim_{n\to\infty}\ \frac{1}{n}\log P_{H_o}\left(T_{n,k} < a_{n,k}(\alpha)\right) = -\alpha$$

$$\lim_{n\to\infty}\ \frac{1}{n}\log P_{H_o}\left(T_{n,k} > b_{n,k}(\alpha)\right) = -\alpha$$

(Existence of $a_{n,k}(\alpha)$, $b_{n,k}(\alpha)$ is easily verified in each case by simple theorems of large deviations (for example [22], [14]). Then the test with critical region :

$$\bigcup_{k=1}^{n}\left\{T_{n,k} < a_{n,k}(\alpha)\right\} \cup \left\{T_{n,k} > b_{n,k}(\alpha)\right\}$$

. is uniformly asymptotically optimal since it has the log-level . coïncide in each case with the likelihood ratio test.

(More details on exact rate of convergence can be found in [15]).

This optimality property of the likelihood ratio test does not remain in the example II of the introduction (about Poisson Process) since, after reducing the problem to exponential variables, (Example 4) the sample size $N(1)$ and the change point $N(\tau)$ are still random variables with fluctuations around their expectations of the same order as the probabilities of the previous critical regions.

b) Optimal boundaries in moments methods

We have seen in I.2 that cusum techniques can be used with various boundaries $g(.)$. The problem of choosing g with optimal properties is a rather tricky problem for n fixed (even large).

In Chernoff sense, optimal boundaries have been computed in [15], for example, in the case of change in mean of a normal model the optimal boundary associated with the cusum test is a parabola, and the optimal boundary associated with the cusum of squares (2^{nd} moment) is solution of :

$$\frac{g^*(t) - 2\alpha}{t} = \log\left[1 + \frac{g^*(t)}{t}\right]$$

(for details see [15]) .

c) Comparison of tests on their separating power.

In the preceding sections (2a, 2b) we did comparisons between tests through their exponential rate of type II error. That method leads rather easily to optimal tests but to compare precisely two procedures which are neither optimal, evaluations of the rates may be rather difficult and it can be useful to turn to this simpler method : when the log-level has a fixed rate, explicit evaluations of the type II error show that for a sequence of test, there are points of the alternative where the type II error is exponentially decreasing, and points where it is not.

We'll give the name of "Range $R_\alpha(\phi_n)$ separated by a sequence of test $\{\phi_n\}$" at the log-level α : the subset of the alternative where the type II error is exponentially decreasing

$$\sup_{\theta \in \Theta_o} \overline{\lim_{n\to\infty}} \frac{1}{n} E_\theta \phi_n \leq -\alpha$$

$$R_\alpha(\phi_n) = \{\theta_1 \in \Theta_1 \,/\, \overline{\lim_{n\to\infty}} \frac{1}{n} \log E_{\theta_1}(1-\phi_n) < 0\}$$

and we can make comparisons between the separating powers of different sequences of tests through their range of separation.

As a matter of fact, knowing that at a point of the alternative, one sequence of test has an exponentially decreasing 2^{nd} type error, and the other not is certainly more informative about the real qualities of the test than an exact comparison between two exponential rates of convergence.

Let us take the example of testing for a change point in the mean of a normal family with known variance :

$$H_o = \{X_1, \ldots, X_n : \text{i.i.d. } \mathcal{N}(\theta_o, 1)\}$$

versus $$H_1^*(\tau, \Delta) = \{X_1, \ldots, X_{k(n)} : \text{i.i.d. } \mathcal{N}(\theta, 1)$$

$$X_{k(n)+1}, \ldots, X_n : \text{i.i.d.} \mathcal{N}(\theta+\Delta, 1), \frac{k(n)}{n} \to \tau\}$$

we have compared the "range" , at the log-level α

$$R = \{(\tau, \Delta), \lim_{n\to\infty} \frac{1}{n} \log E_{H_1^*(\tau,\Delta)} \phi_n < 0\} \text{ for :}$$

- Cusum of the recursive residuals w_i : crossing the linear boundary : $g(t) = \sqrt{\alpha}\, t + \frac{\sqrt{\alpha}}{2}$:

$$\mathcal{D}_1 = \{\exists k = 1, \ldots, n , \frac{|\sum_1^k w_i|}{n} > \sqrt{\alpha} \cdot \frac{k}{n} + \frac{\sqrt{\alpha}}{2}\}$$

resp. the optimal boundary $g^*(t) = \sqrt{2\alpha t}$: $\mathcal{D}_1'$.

- Cusum of the squares of the recursive residuals : crossing the linear boundary $g(t) = t + c(\alpha)$: $\mathcal{D}_2$.

- Likelihood ratio test : the constant boundary α gives the log-level α and using the recursive residuals, this test admits the critical region :

$$\mathcal{D}_3 = \{\exists k = 1, \ldots, n-1, \frac{k}{2(n-k)} \left(\sum_{k+1}^{n} \frac{w_i}{\sqrt{i(i-1)}}\right)^2 \geq \alpha\}$$

A detailed investigation is given in [15] ; it leads to the respective ranges R_1, R'_1, R_2, R_3 which can be represented by table 1 ; and its conclusions can be summarized as follows :

. the improvement of $\mathcal{D}_1'$ with respect to $\mathcal{D}_1$ is especially appreciable in the neighbourhood of $\tau = 0$. The "ranges" R_1 and R'_1 are not symmetric : the detection of a change is more sensitive for a little value τ than for the symmetric one $1-\tau$; the best "range" is observed at $\tau = e^{-1}$.

. the range R_2 is exactly the same for the test based on the whole variance which has critical region :

$$\mathcal{D}_2' = \left\{ \frac{\sum_{i=2}^{n} w_i^2}{n-1} > 1 + c(\alpha) \right\}$$

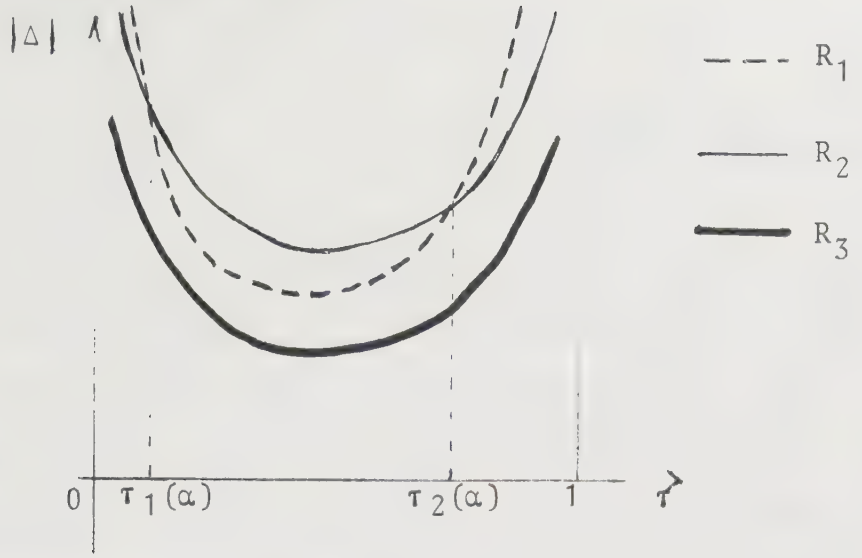
This remark induces the question : wouldn't be interesting to save up a long computing time by using the whole variance test instead of the cusum of squares for a negligible loss. A partial answer will be given with simulation results.

. It is clear that the LR Test remains significantly better than others.

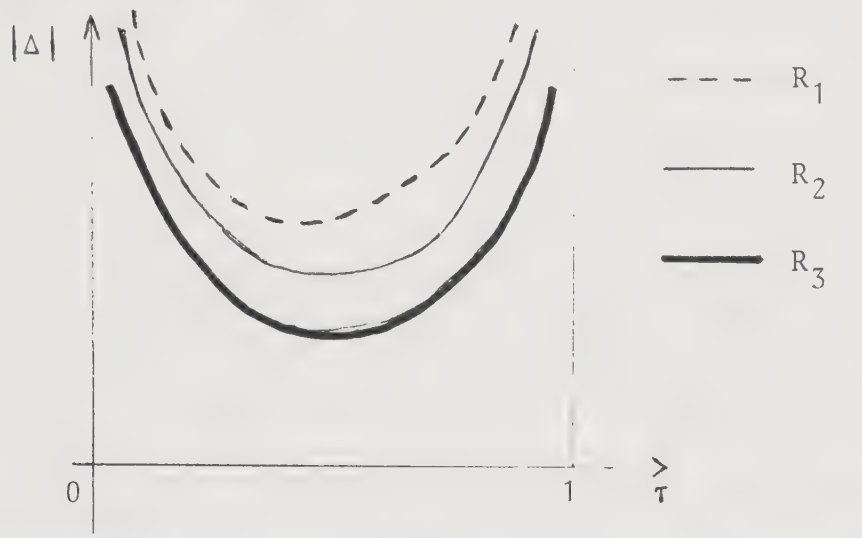
. For small values of the log-level α, the range R_1 of the cusum test is strictly better than the cusum of squares one : R_2 for values of τ belonging to an interval $[\tau_1(\alpha), \tau_2(\alpha)]$. It will be useful to notice that because of the preceding remarks, $\tau_1(\alpha)$ is closer to 0 than $\tau_2(\alpha)$ to 1.

. The comparison for larger values of α makes little sens

. for small values of α



. for large values of α



. for small values of α and a parabolic boundary, the range is improved at small values of τ :

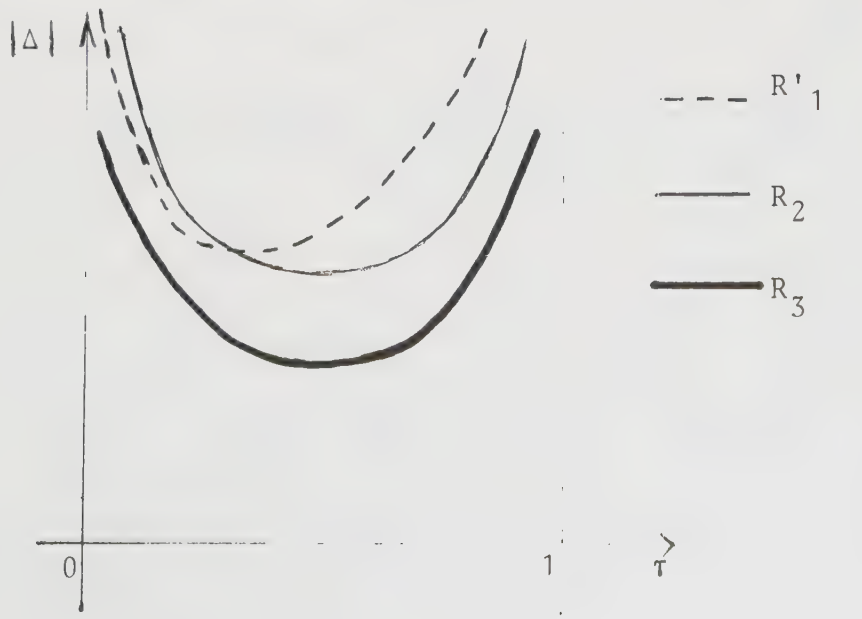


TABLE 1

3 - MONTE-CARLO'S COMPUTATIONS

We have computed the power of the previous tests (cusum, cusum of squares and likelihood ratio tests) in the model described in part II.2.c). Before computing the power, we evaluated the level of the tests again with Monte-Carlo methods (an approximation of the exact level, using limit theorems gives a precision of about $n^{-1/2}$, though the precision of Monte-Carlo methods is of order $N^{-1/2}$ where N is the number of iterations ; approximating such very little quantities as test levels, we observed that the first approximation of a relative error of about $\frac{1}{2}$).

We have considered two values of n : 30 (table 3) and 200 (table 4) and different values of the change point k and of the deviation Δ . All the levels and powers have been computed with N=1 000 iterations. We get the powers of :

- on the first line, the cusum test ($\mathscr{D}_1$)
- on the second line, the cusum of squares test($\mathscr{D}_2$)
- on the third line, the whole variance test ($\mathscr{D}_2'$)
- on the fourth line, the likelihood ratio test ($\mathscr{D}_3$).

These two tables suggest the following remarks :

TABLE 2

COMPUTED POWER FOR N=30 AT LEVEL 0.05

	k=2	k=5	k=10	k=15	k=20	k=25	k=28
Δ=0.5	0.08* 0.06 0.06 0.09	0.14* 0.08 0.07 0.15	0.18* 0.09 0.08 0.22	0.16* 0.09 0.09 0.24	0.10* 0.08 0.08 0.22	0.07* 0.07 0.07 0.15	0.06 0.06 0.06 0.09
Δ=1	0.20* 0.10 0.09 0.20	0.42* 0.17 0.15 0.42	0.52* 0.24 0.23 0.65	0.47* 0.25 0.26 0.72	0.29* 0.22 0.23 0.65	0.12 0.14 0.15 0.42	0.07 0.08 0.09 0.20
Δ=1.5	0.37* 0.17 0.15 0.39	0.72* 0.36 0.32 0.75	0.85* 0.56 0.55 0.94	0.80* 0.60 0.61 0.97	0.57* 0.51 0.55 0.94	0.23 0.29 0.32 0.75	0.08 0.13 0.15 0.39
Δ=2	0.58* 0.32 0.26 0.63	0.92* 0.64 0.59 0.95	0.98* 0.87 0.86 0.997	0.96* 0.91 0.92 0.999	0.82 0.84 0.86 0.997	0.36 0.54 0.59 0.95	0.10 0.22 0.26 0.63
Δ=2.5	0.76* 0.47 0.41 0.82	0.993* 0.87 0.85 0.996	1* 0.99 0.99 1.	0.998* 0.993 0.995 1.	0.96 0.98 0.992 1.	0.52 0.81 0.85 0.996	0.13 0.36 0.41 0.82
Δ=3	0.90* 0.69 0.60 0.95	1.* 0.98 0.97 1.	1. 1. 1. 1.	1. 1. 1. 1.	0.997 1. 1. 1.	0.68 0.96 0.97 1.	0.17 0.56 0.60 0.95
Δ=4	0.993* 0.95 0.90 0.999	1. 1. 1. 1.	1. 1. 1. 1.	1. 1. 1. 1.	1. 1. 1. 1.	0.90 0.999 1. 1.	0.28 0.88 0.90 0.999

TABLE 3

COMPUTED POWER FOR N=200 AT LEVEL 0.05

	k=5	k=10	k=20	k=50	k=100	k=150	k=180	k=190	k=195
=0.5	0.11* 0.06 0.06 0.13	0.20* 0.07 0.07 0.21	0.40* 0.08 0.08 0.41	0.67* 0.13 0.13 0.74	0.64* 0.15 0.16 0.86	0.28* 0.11 0.13 0.74	0.08* 0.06 0.08 0.41	0.06 0.05 0.07 0.21	0.05 0.05 0.06 0.13
=1	0.28* 0.09 0.09 0.38	0.60* 0.13 0.13 0.70	0.93* 0.26 0.25 0.95	0.998* 0.57 0.57 1.	0.998* 0.71 0.76 1.	0.82* 0.49 0.57 1.	0.21 0.17 0.25 0.95	0.08 0.08 0.13 0.70	0.06 0.06 0.09 0.38
=1.5	0.56* 0.16 0.15 0.74	0.93* 0.34 0.30 0.97	0.999* 0.67 0.61 1.	1.* 0.97 0.96 1.	1.* 0.99 0.99 1.	0.99* 0.94 0.95 1.	0.40 0.56 0.61 1.	0.14 0.24 0.30 0.97	0.07 0.12 0.15 0.74
=2	0.81* 0.31 0.26 0.96	0.997* 0.66 0.57 1.	1.* 0.96 0.93 1.	1. 1. 1. 1.	1. 1. 1. 1.	1. 1. 1. 1.	0.69 0.89 0.93 1.	0.23 0.46 0.57 1.	0.09 0.18 0.26 0.96
=2.5	0.95* 0.53 0.43 0.998	1.* 0.93 0.86 1.	1. 1. 1. 1.	1. 1. 1. 1.	1. 1. 1. 1.	1. 1. 1. 1.	0.89 0.994 1. 1.	0.33 0.78 0.86 1.	0.11 0.33 0.43 0.998
=3	0.994* 0.79 0.66 1.	1.* 0.998 0.98 1.	1. 1. 1. 1.	1. 1. 1. 1.	1. 1. 1. 1.	1. 1. 1. 1.	0.98 1. 1. 1.	0.48 0.95 0.97 1.	0.15 0.55 0.66 1.

. The cusum test ($\mathscr{D}_1$) detects with more precision, a change point at the beginning (small τ). On the other hand, the power is quite symmmetric for the cusum of squares test ($\mathscr{D}_2$).

As expected, there is no uniform comparison between the cusum of square test and the whole variance one : it is however interisting to note that the former is better for small values of τ (underlined values in tables 2 and 3) and the latter for large ones.

. The differences between the cusum ($\mathscr{D}_1$) and the cusum of squares ($\mathscr{D}_2$) are generally more emphasized : the former is more powerful for small values of τ and Δ (starred values in tables 2 and 3).

. The L.R. test ($\mathscr{D}_3$) is obviously more powerful than all the others , a significant difference only appearing at very few small values of τ and Δ .

Confrontation with asymptotic results.

The comparisons are essentially coherent : nearness of the test $\mathscr{D}_2$ and $\mathscr{D}'_2$, maximal power obtained at τ near e^{-1} for the $\mathscr{D}_1$, optimality of $\mathscr{D}_3$.

The only contradiction is the fact that we didn't observe a better behaviour for the test $\mathscr{D}_2$ (with respect to $\mathscr{D}_1$) at small values of τ : it could be explained by the essential asymmetry of the cusum test power, and the fact that τ_1 is much closer to 0 than τ_2 to 1 .

Furthermore, it is very interesting to note that the curves demarcating the "ranges" are quite the same as those joining the points of equal power.

4 - MULTI-INDICE CONVERGENCE - MAIN THEOREMS.

The purpose of this section is to emphasize the advantage of a generalization to partial sums processes of classical limit theorems in statistic.

We shall first give a general invariance theorem that can be used to generalize a large class of procedures for comparison between populations(or goodness of fit) to change point models (especially, it may be useful to construct on the same technique, tests based on classical M-estimators). Then, we'll give examples of applications of this theorem needed in the previous sections

Theorem 1 :

Under the following assumptions

. Λ is a compact interval of $\mathbb{R}$, including 0

. $\xi_{n,1}(.)$, $\xi_{n,2}(.),\ldots,$ $\xi_{n,n}(.)$ are random variables on to $\mathscr{C}(\Lambda)$, independant, identically distributed, zero mean, and such that they verify one of the two moments conditions (M1) or (M2) ($\xi_{n,i}(0)$ being null for all indices).

(M1) There exist

- a continuous, non decreasing function G on Λ
- a constant α greater than 1

$$\forall i,\ \forall n,\ \forall \lambda_1 < \lambda_2,\quad E\,|\xi_{n,i}(\lambda_2) - \xi_{n,i}(\lambda_1)|^2 \leq \frac{[G(\lambda_2)-G(\lambda_1)]^{\alpha}}{n}$$

(M2) There exist :

- a continuous, non decreasing function G on Λ

- a constant $\alpha > 1$ such that :

$\forall\, i, \forall\, n\,, \forall\, \lambda_1 < \lambda_2 < \lambda_3$:

$$E\,(\xi_{n,i}(\lambda_2) - \xi_{n,i}(\lambda_1))^2 \leq \frac{G(\lambda_2) - G(\lambda_1)}{n}$$

$$E\,(\xi_{n,i}(\lambda_2) - \xi_{n,i}(\lambda_1))^2\,(\xi_{n,i}(\lambda_3) - \xi_{n,i}(\lambda_2))^2 \leq \frac{(G(\lambda_2)-G(\lambda_1))^{\alpha}}{n}$$

The partial sum process defined by :

$$[0,1] \times \Lambda \;\rightarrow\; \mathbb{R}$$

$$(t,\lambda) \qquad Y_n(t,\lambda) = \sum_{i=1}^{[nt]} \xi_{n,i}(\lambda) + (nt-[nt])\,\xi_{n,[nt]+1}$$

is such that whenever the finite dimensional distributions of the mono-indice process $Y_n(1,.)$ converge to those of a gaussian process $L(.)$, then the double indiced process $Y_n(.,.)$ converges in $\mathscr{C}([0,1] \times \Lambda)$ to the Gaussian process $Y(t,\lambda)$, zero mean and with covariance $E\,Y(t,\lambda)\,Y(t',\lambda') = t \wedge t'\ \mathrm{cov}\,(L(\lambda), L(\lambda'))$.

Proof of the theorem :

1. t fixed, convergence of the process $Y_n(t,.)$ in $\mathscr{C}(\Lambda)$ to $\sqrt{t}\,L(.)$:

- convergence of the finite dimensional distribution are immediate
- as for the tightness, it is easy to remark that for $Y_n(t,.)$ conditions of theorem 12.3 (respectively 12.84) of Billingsley [4] are verified if (M1) (respectively M2) is true.

2. Tightness of the bidimensional process.

In spite of the multi indice type of the problem we'll not use here the results of Bickel-Wichura, because, for the parameter t, mor refined conditions are needed and particularly the structure of independant increaments process, as can be pointed in the following lemma.

Lemma 1 : Let $\eta_1(.),\ldots,\eta_n(.)$ random functions of $\mathscr{C}(\Lambda)$, independant, centered, with finite variances $\sigma_i^2(.)$.

Let $S_i(\lambda) = \eta_1(\lambda) + \ldots + \eta_i(\lambda)$

and $(s_m(.))$ be a sequence of strictly positive function defined on Λ .

$$\left[1 - \sup_{0\le k\le m} P\left\{\exists\, \lambda\in\Lambda,\ |S_m(\lambda)-S_k(\lambda)| > \frac{A}{2}\, s_m(\lambda)\right\}\right] \times$$

$$P\left(\exists i = 0,\ldots,m,\ \exists\, \lambda\in\Lambda,\ |S_i(\lambda)| > A\, s_m(\lambda)\right) \le$$

$$P\left\{\exists\, \lambda\in\Lambda,\ |S_m(\lambda)| \ge \frac{A}{2}\, s_m(\lambda)\right\}$$

Proof of the lemma :

Put $E_k = \{S_j(\lambda) \le A\, s_m(\lambda),\ \forall\, \lambda\in\Lambda,\ \forall j = 1,\ldots,k-1\}$

$$\cap\ \{\exists\, \lambda\in\Lambda\ |S_k(\lambda)| > A\, s_m(\lambda)\}$$

then :

$$P\left\{\exists k = 0,\ldots,m,\ \exists\, \lambda\in\Lambda,\ |S_k(\lambda)| > A\, s_m(\lambda)\right\}$$

$$\le P\left\{\forall\, \lambda\in\Lambda,\ |S_m(\lambda)| > \frac{A}{2}\, s_m(\lambda)\right\} +$$

$$\sum_{k=1}^{m} P\left(E_k \cap \forall\, \lambda\in\Lambda,\ |S_m(\lambda)| < \frac{A}{2}\, s_m(\lambda)\right\}$$

and :

$$P\ (E_k \cap \forall\lambda\in\Lambda,\ |S_m(\lambda)| < \frac{A}{2}\ s_m(\lambda)\} \leq P(E_k).P(\exists\lambda\in\Lambda,\ |S_m(\lambda)-S_k(\lambda)|>\frac{A}{2}.s_m(\lambda))$$

Replacing in the first inequality, we obtain :

$$P(\sup_{k=1,\ldots,m}\ \sup_{\lambda\in\Lambda} |\frac{S_k(\lambda)}{s_m(\lambda)}| > A)$$

$$\leq P(\sup_{\lambda\in\Lambda} \frac{|S_m(\lambda)|}{s_m(\lambda)} > \frac{A}{2}) + \sup_{k=1,\ldots,m-1} P(\sup_{\lambda\in\Lambda}\frac{|S_m(\lambda)-S_k(\lambda)|}{s_m(\lambda)} > \frac{A}{2}) \times \sum_{k=1}^{m-1} P(E_k)$$

$$\leq P(\sup_{\lambda\in\Lambda} \frac{|S_m(\lambda)|}{s_m(\lambda)} > \frac{A}{2}) + \sup_{k=1,\ldots,m-1} P(\sup_{\lambda\in\Lambda} \frac{|S_m(\lambda)-S_k(\lambda)|}{s_m(\lambda)} > \frac{A}{2}\}$$

$$\times\ P\ (\sup_{k=1,\ldots,m}\ \sup_{\lambda\in\Lambda} \frac{|S_k(\lambda)|}{s_m(\lambda)} > A)$$

which completes the proof of the lemma.

Now, for proving tightness we must show the two inequalities:

$\forall\ \varepsilon > 0,\ \forall\ A > 0,\ \exists\ \delta_o\ (\varepsilon,A),\ \exists\ n_o\ (\varepsilon,A,\delta),\ \forall\ \delta \leq \delta_o,\ \forall\ n \geq n_o$:

$$*\qquad P\ (\sup_{\lambda\in\Lambda}\ \sup_{|t_1-t_2|<\delta} |Y_n\ (t_2,\lambda) - Y_n\ (t_1,\lambda)| > A) \leq \varepsilon$$

and

$$**\qquad P(\sup_{t\in[0,1]}\ \sup_{|\lambda_1-\lambda_2|<\delta} |Y_n\ (t,\lambda_1) - Y_n\ (t,\lambda_2)| > A) \leq \varepsilon$$

Proof of inequality (*).

Applying lemma 1 we can replace (*) by

$$(*1)\qquad \sup_{k=1,\ldots,n\delta}\ P\ (\sup_{\lambda\in\Lambda} |\sum_{i=1}^{k} \xi_{n,i}\ (\lambda)| > \frac{A}{4}\} \leq \frac{1}{2}$$

$$(*2)\qquad P\ (\sup_{\lambda\in\Lambda} |\sum_{i=1}^{n\delta} \xi_{n,i}\ (\lambda)| > \frac{A}{4}\} \leq \varepsilon\delta$$

(*1) : We follow the method of Billingsley [4], dividing the event by using variables :

$$U_n(\ell) = \sum_{j=1}^{k} \xi_{n,j}\left(\frac{\ell}{m}\right) - \xi_{n,j}\left(\frac{\ell-1}{m}\right), \quad \ell \in \{m(\lambda_{min}), m\lambda_{max}\}$$

They verify the conditions of theorem (12.2) of Billingsley (under M1) or theorem (12.1) (under M2) and thus

$$\text{if (M1)}, \quad P\Big(\sup_{1\le p\le(\lambda_{max}-\lambda_{min})m} \Big|\sum_{\ell=m\,\lambda_{min}}^{p+m\,\lambda_{min}} U_n(\ell)\Big| > \frac{A}{4}\Big\} \le \frac{C}{A^2}\frac{k}{n}\left[G(\lambda_{max}) - G(\lambda_{min})\right]^{\alpha}$$

The right side being free of m, m may tend to infinity :

$$\forall k \le n\delta, \quad P\Big(\sup_{\lambda\in\Lambda}\Big|\sum_{j=1}^{k}\xi_{n,j}(\lambda)\Big| > \frac{A}{4}\Big\} \le \frac{C}{A^2}\,\delta\left[G(\lambda_{max}) - G(\lambda_{min})\right)^{\alpha}$$

$$\text{if (M2)}, \quad P\Big(\sup_{p}\Big|\sum_{\ell=m\,\lambda_{min}}^{p+m\,\lambda_{min}} U_n(\ell)\Big| > \frac{A}{4}\Big) \le \frac{C}{A^4}\left[G(\lambda_{max}) - G(\lambda_{min})\right]^{\alpha}\left(\frac{k}{n}\right.$$

$$+ P\Big\{\max_{p}|U_n(p)| > \frac{A}{8}\Big)$$

$$\le \frac{C\delta}{A^4} + \sum_{p} P\Big\{\Big|\sum_{1}^{k}\xi_{n,j}\left(\frac{\ell}{m}\right) - \xi_{n,j}\left(\frac{\ell-1}{m}\right)\Big| > \frac{A}{8}\Big\}$$

$$\le \frac{C\delta}{A^4} + \frac{C\delta}{A^2}\left[G(\lambda_{max}) - G(\lambda_{min})\right]$$

m tending to infinity, we obtain the result.

(*2) : Using convergence of $Y_n(\delta,.)$ to $\sqrt{\delta}\,L(.)$ in $\mathscr{C}(\Lambda)$ and Tchebichev Inequality we obtain :

$$P\Big(\sup_{\lambda\in\Lambda}|Y_n(\delta,\lambda)| > \frac{A}{4}\Big) \le 2\delta^2\left[\left(\frac{4}{A}\right)^4 E\sup_{\lambda\in\Lambda}|L(\lambda)|^4\right. \quad \text{for}$$

$n \ge N(\delta, A)$, the right side being finite since L(.) is a gaussian

process of $\mathscr{C}(\Lambda)$ ($\sup_{\lambda\in\Lambda} L(\lambda) < +\infty$ a.e) : see Fernique [30].

. Inequality (**) can be shown in the same way. Using the lemma, it remains to get

$$(**1)\quad \sup_{k=1,\ldots,n} P\Big\{ \sup_{|\lambda_1-\lambda_2|<\delta} \Big| \sum_{j=1}^{k} \xi_{n,j}(\lambda_2) - \sum_{j=1}^{k} \xi_{n,j}(\lambda_1)\Big| > \frac{A}{2}\Big\} \le 1-\eta$$

$$(**2)\quad P\Big(\sup_{|\lambda_2-\lambda_1|<\delta} \Big| \sum_{j=1}^{n} \xi_{n,j}(\lambda_2) - \sum_{j=1}^{n} \xi_{n,j}(\lambda_1)\Big| > \frac{A}{2}\Big\} \le \frac{\varepsilon}{2}$$

This last inequality is a consequence of the tightness of the process $Y_n(1,.)$.

For the first inequality, dividing $[0,1]$ in intervals of length δ and centers t_i^* , we obtain :

$$\sup_{k=1,\ldots,n} P\Big(\sup_{|\lambda_1-\lambda_2|<\delta} \Big| \sum_{j=1}^{k} \xi_{n,j}(\lambda_2) - \sum_{j=1}^{k} \xi_{n,j}(\lambda_1)\Big| > \frac{A}{2}\Big\}$$

$$= \sup_{t_1^*,t_2^*,\ldots,t_\ell^*} \sup_{k,\left|\frac{k}{n}-t_i^*\right|<\frac{\delta}{2}} P\Big(\sup_{|\lambda_1-\lambda_2|<\delta} |Y_n(\tfrac{k}{n},\lambda_1) - Y_n(\tfrac{k}{n},\lambda_2)| > \frac{A}{2}\Big\}$$

$$\le 2P\Big\{ \sup_{|t_1-t_2|<\delta} \sup_{\lambda} |Y_n(t_1,\lambda)-Y_n(t_2,\lambda)| > \frac{A}{6}\Big\}$$

$$+ \sup_{t_1^*,\ldots,t_\ell^*} P\Big(\sup_{|\lambda_1-\lambda_2|<\delta} |Y_n(t_i^*,\lambda_1)-Y_n(t_i^*,\lambda_2)| > \frac{A}{6}\Big)$$

$\le 2\varepsilon + (1-2\varepsilon-\eta)$, for n sufficiently large.

2ε coming from inequality (*)

$1-2\varepsilon-\eta$ obtained by using, for fixed δ and t_i^* , the convergence of $Y_n(t_i^*,.)$ in $\mathscr{C}(\Lambda)$.

It remains to take (for instance) $\eta = \sqrt{\varepsilon}$ to complete the proof.

1. Consequence : Empirical process - Kolmogorov-Smirnov test.

Using the previous theorem, we can obtain the classical invariance principle on ecdf [32]:

Let $U_1,\ldots,U_n$ be n random variables, independant, uniformly distributed on $[0,1] = \Lambda$.

Put

$$\xi_{n,i}(\lambda) = \frac{1}{\sqrt{n}} \left(1_{\{U_i \le \frac{\lambda}{1-\epsilon_n}\}} - \frac{1}{\epsilon_n}\left(1 - \frac{\lambda}{U_i}\right) 1_{\{\lambda \le U_i \le \frac{\lambda}{1-\epsilon_n}\}} + \lambda . \frac{\log(1-\epsilon_n)}{\epsilon_n}\right)$$

then the processes :

$$Y_n(t,\lambda) = \sum_{i=1}^{[nt]} \xi_{n,i}(\lambda) + (nt-[nt])\, \xi_{n,[nt]+1}$$

and $Y'_n(t,\lambda) = Y_n(t,\lambda) - t\, Y_n(1,\lambda)$

converge in $\mathscr{C}([0,1]^2)$ respectively to the process

$B(.,.)$ and $B'(.,.)$ where

$B(t,\lambda) = W(t,\lambda) - \lambda\, W(t,1)$

$B'(t,\lambda) = W(t,\lambda) - \lambda\, W(t,1) - t\, W(1,\lambda) + \lambda t\, W(1,1)$ is the Kiefer process, $W(.,.)$ being the Wiener process on $[0,1]^2$.

The result if obviously obtained by noticing that the $\xi_{n,i}(.)$ verify condition (M2) with $G(\lambda)\equiv\lambda, \alpha=2$ and give for $Y_n(.,.)$ the same limit finite dimensional distributions, as $\xi_{n,i}(\lambda) = \frac{1}{\sqrt{n}} \left(1_{U_i<\lambda} - \lambda\right)$.

2. Consequence : Likelihood process.

Theorem 2 :

Let $X_1,\ldots,X_n$ be a n-sample of the real distribution function F_θ ($\theta \in \Theta$, where Θ is an interval included in $\mathbb{R}$). $\{F_\theta\}_{\theta\in\Theta}$ is supposed to be dominated by a measure μ and the associated densities verify the regularity asumptions :

A1 The function $\theta \to \log f(x,\theta)$ is twice continuously differentiable for all θ in a neighbourhood of θ_o : $V(\theta_o)$, for almost every x (with respect to measure μ)

A2 $$I(\theta) = E_\theta \left(\frac{f'(x,\theta)}{f(x,\theta)}\right)^2 < +\infty \ , \ \theta\in V(\theta_o), \ (f'(x,\theta) \text{ denote } \frac{\partial}{\partial\theta}))$$

A3 $$E_\theta \left|\frac{\partial^2}{\partial\theta^2} (\log f(x,\theta))\right| < +\infty \ , \ \theta\in V(\theta_o)$$

A4 Let $$w(x,\delta,\theta) = \sup_{|\theta'-\theta|<\delta} \left| \frac{\partial^2}{\partial\theta^2} \log f(x,\theta') - \frac{\partial^2}{\partial\theta^2} \log f(x,\theta)\right|$$

$$\exists\, \delta_o \ / \quad E_{\theta_o} w(x,\delta_o,\theta_o) < +\infty$$

When Λ is a compact neighbourhood of 0, the likelihood process $Z_{n,\theta_o}(.,.)$ defined by :

$$Z_{n,\theta_o}(t,\lambda) = \sum_{i=1}^{[nt]} \log \frac{f\left(X_i,\theta_o + \frac{\lambda}{\sqrt{n}}\right)}{f(X_i,\theta_o)} + (nt-[nt])\, x$$

$$\log \frac{f\left(X_{[nt]+1},\, \theta_o + \frac{\lambda}{\sqrt{n}}\right)}{f(X_{[nt]+1},\theta_o)}$$

converges, if θ_o is true, in $\mathscr{C}([0,1] \times \Lambda)$ to the gaussian process $Z(.,.)$ defined by

$$Z(t,\lambda) = \lambda \sqrt{I(\theta_o)}\, W(t) - \frac{\lambda^2 I(\theta_o)}{2} t$$

Proof : Classical techniques easily show that regularity conditions allow to assure that the

$$\xi_{n,i}(\lambda) = \text{Log} \frac{f\left(X_i, \theta_o + \frac{\lambda}{\sqrt{n}}\right)}{f(X_i,\theta_o)} - E_{\theta_o} \log \frac{f\left(X_i, \theta_o + \frac{\lambda}{\sqrt{n}}\right)}{f(X_i,\theta_o)}$$

are sum of

$$\xi'_{n,i}(\lambda) = \frac{\lambda}{\sqrt{n}} (\log f)'(X_i,\theta_o) + \frac{\lambda^2}{2n} (\log f)''(X_i,\theta_o) + \frac{\lambda^2}{2n} I(\theta_o) :$$

and $R_{n,i}(\lambda)$, with $n\, R_{n,i}(\lambda)$ tending to 0 uniformly for $\lambda \in \Lambda$, (verifying M1) when n tends to infinity.

The functionnal :

$$G : \mathscr{C}([0,1]) \times \Theta \rightarrow \mathscr{C}[0,1]$$

$$f(t,\theta) \rightarrow G(f)(t) = 2t \sup_{\theta} f(t,\theta)$$

being continuous we have the

Corollary 1 :

Under asumptions A1 ... A4, the one dimensional sup-process : $\{Z_{n,\theta_o}(t) \ , \ t \in [0,1]\}$ converges when n goes to infinit to $\{2t \sup_{\lambda\in\Lambda} Z(t,\lambda), \ t \in [0,1]\}$

where $$Z_{n,\theta_o}(t) = 2t \sup_{\lambda\in\Lambda} Z_{n,\theta_o}(t,\lambda)$$

Our purpose, now is to obtain a similar result for the process $\{Z^*_{n,\theta_o}(t), \ t \in [0,1]\}$ defined in corollary 2. The difference between these two processes essentially lies in compacity of Λ .

Lemma 2: For every positive η there exists n_o such that for n greater than n_o :

the function : $\theta \to \frac{1}{n} \sum_{i=[nt]}^{n} \log f(X_i, \theta)$ admits a local maximum θ_t^* in the interval $[\theta_o-\eta, \theta_o+\eta]$, for all t in $[0,1-\varepsilon]$.

Proof : By strong Law of Large Number we get : The 3-dimensional vector :

$$\left(\frac{1}{j} \sum_{i=1}^{j} \log f(X_i,\theta_o-\eta), \frac{1}{j} \sum_{1}^{j} \log f(X_i,\theta_o+\eta), \frac{1}{j} \sum_{i}^{j} \log f(X_i,\theta_o)\right)$$

converges a.e to $(E_{\theta_o} \log f(X_i, \theta_o-\eta), E_{\theta_o} \log f(X_i,\theta_o+\eta), E_{\theta_o} \log f(X_i,\theta_o))$.

Hence, for every δ , there exist $n_1(\delta)$, such that we have a.e., for $j > n_1$

$$\left|\frac{1}{j} \sum_{i=1}^{j} \log f(X_i,\theta_o-\eta) - E_{\theta_o} \operatorname{Log} f(X, \theta_o-\eta)\right| \le \delta$$

$$\left|\frac{1}{j} \sum_{i=1}^{j} \log f(X_i,\theta_o+\eta) - E_{\theta_o} \log f(X, \theta_o+\eta)\right| \le \delta$$

$$\left|\frac{1}{j} \sum_{i=1}^{j} \log f(X_i,\theta_o) - E_{\theta_o} \log f(X, \theta_o)\right| \le \delta$$

We just have to take $\delta < \frac{1}{2} \inf \{ E_{\theta_o} \log \frac{f(X,\theta_o+\eta)}{f(X,\theta_o)}, E_{\theta_o} \log \frac{f(X,\theta_o-\eta)}{f(X,\theta_o)} \}$

and $n_o = \frac{n_1(\delta)}{\varepsilon}$ to derive the proof of the lemma.

* Notice that the result is false replacing $[0,1-\varepsilon]$ by $[0,1]$.

Corollary 2 : Under assumptions A1 ... A4 and if for each k in $\{1,\ldots,n\}$ equation (A) :

$$\sum_{i=n-k+1}^{n} \frac{f'}{f}(X_i,\theta) = 0$$

has an unique solution $\hat{\theta}^*_k$ then : the one dimensional proces

$\{Z^*_{n,\theta_o}(t)\ ,\ t \in [0,1]\}$ converges to $\{w^2(t),\ t \in [0,1]\}$ when n goes to infinity,

where $Z^*_{n,\theta_o}(.)$ is the polygonal line constructed on the points

$$\{(0,0),\ (\frac{k}{n},\ \frac{2k}{n} \sum_{i=n-k+1}^{n} \log \frac{f(X_i,\hat{\theta}^*_k)}{f(X_i,\theta_o)})\quad k = \{1,\ldots,n\}\ .$$

Proof : Corollary 1 and Lemma 2 imply that the traces of the two processes :

$$Z_{n,\theta_o}(t) \quad \text{et} \quad Z^*_{n,\theta_o}(t) = \sup_{\lambda \in \mathbb{R}} 2t\, Z_{n,\theta_o}(t,\lambda)$$

on $\mathscr{C}[\delta,1]$ are equal on a set of probability greater than $1-\frac{\varepsilon}{2}$ for n greater than $N(\varepsilon,\delta)$. For, using a Taylor expansion (as in p. 44) and assumption A4, it may be seen that lemma 2 implies that $\sqrt{n}\ (\hat{\theta}_k-\theta_o)$ lies in a compact set uniformly for $k \in n[\delta,1]$ with a probability greater than $1-\varepsilon/4$).

Furthermore the traces of the "limit processes"

$Z^*_{\theta_o}(t) = W^2(t)$ and

$$Z_{\theta_o}(t) = 2t \sup_{\lambda \in \Lambda} \lambda \sqrt{I(\theta_o)}\, W(t) - \lambda^2 \frac{I(\theta_o)t}{2}$$

on $([\delta,1])$ also are identical on a set of probability greater than $1-\frac{\varepsilon}{2}$ (for a compact set Λ large enough).

Thus ε being arbitrary fixed we deduce that the trace of $Z^*_{n,\theta_o}(.)$ converges to that of $Z^*_{\theta_o}(.)$ in $\mathcal{C}([\delta,1])$ for all δ $(0 < \delta < 1)$.

To complete the proof it remains to show that the sequence $Z^*_{n,\theta_o}(.)$ is tight in a neighbourhood of 0, i.e.

$$\forall \varepsilon > 0,\ \forall A > 0,\ \exists \delta > 0,\ \exists N(\varepsilon,A,\delta) \ /\ \forall n > N(\varepsilon,A,\delta)\ ,$$

$$P\left(\sup_{k=1,\ldots,n\delta} \left| 2\frac{k}{n} \sum_{i=1}^{k} \log \frac{f(X_i,\hat{\theta}_k)}{f(X_i,\theta_o)} \right| > A \right\} < \varepsilon$$

First, let us remark that nothing is changed if we take the intersection with the event :

$$\left\{ \sup_{k=1,\ldots,n\delta} |\hat{\theta}_n - \theta_o| > 0 \right\} .$$

A second order Taylor expansion gives :

$$2\frac{k}{n} \sum_{i=1}^{k} \log \frac{f(X_i,\hat{\theta}_k)}{f(X_i,\theta_o)} = -2\frac{k}{n}(\hat{\theta}_k-\theta_o) \sum_{i=1}^{k} \frac{f'}{f}(X_i,\hat{\theta}_k) - \frac{k}{n}(\hat{\theta}_k-\theta_o)^2 \sum_{i=1}^{k} (\log f)''(X_i,\tilde{\theta}_k)$$

$$\tilde{\theta}_k \in (\theta_o,\hat{\theta}_k)$$

and

$$\sum_{i=1}^{k} \frac{f'}{f}(X_i,\hat{\theta}_k) = 0 = \sum_{i=1}^{k} \frac{f'}{f}(X_i,\theta_o) + (\hat{\theta}_k-\theta_o) \sum_{i=1}^{k} (\log f)''(X_i,\theta^*_k)$$

$$\theta^*_k \in (\theta_o,\hat{\theta}_k)$$

and thus

$$2\frac{k}{n} \sum_{i=1}^{k} \log \frac{f(X_i,\hat{\theta}_k)}{f(X_i,\theta_o)} = \frac{\left[\frac{1}{\sqrt{n}} \sum_{i=1}^{k} f'(X_i,\theta_o)\right]^2 \left[\frac{1}{k} \sum_{i=1}^{k} (\log f)''(X_i,\tilde{\theta}_k)\right]}{\left[\frac{1}{k} \sum_{i=1}^{k} (\log f)''(X_i,\theta^*_k)\right]^2}$$

By strong law of large numbers we have

$$\forall \varepsilon > 0, \quad \forall \eta > 0, \quad \exists\, K(\varepsilon,\eta) \;/$$

$$P\{\exists k > K,\ |\hat{\theta}_k - \theta_o| > \eta\} < \frac{\varepsilon}{12} \quad \text{and}$$

$$P\left\{\exists k > K,\ \left|\frac{1}{k}\sum_{1}^{k} (\log f)''(X_i,\theta_o) + I(\theta_o)\right| > \frac{I(\theta_o)}{4}\right\} < \frac{\varepsilon}{12}$$

Furthermore A4 implies : there exist η^* /

$$\exists\, K'(\varepsilon,\eta^*) \;/\; P\left\{\exists k > K',\ \left|\frac{1}{k}\sum_{i=1}^{k} w(X_i,\theta_o,\eta^*)\right| > \frac{I(\theta_o)}{4}\right\} < \frac{\varepsilon}{12}$$

Thus for $K(\varepsilon) = \text{Sup}\,(K(\varepsilon,\eta^*), K'(\varepsilon,\eta^*))$

$$P\left(\exists k > K(\varepsilon),\ \left|\frac{1}{k}\sum_{i=1}^{k} (\log f)''(X_i,\theta_k^*)\right| < \frac{I(\theta_o)}{2}\right\} \le \frac{\varepsilon}{4} \qquad (T_1)$$

$$\text{and } P\left(\exists k > K(\varepsilon),\left|\frac{1}{k}\sum_{i=1}^{k} (\log f)''(X_i,\tilde{\theta}_k)\right| > \frac{3I(\theta_o)}{2}\right\} \le \frac{\varepsilon}{4} \qquad (T_2)$$

Furthermore assumptions A1...A4 are largely enough to ensur convergence of the partial sums process : $\frac{1}{\sqrt{I(\theta_o)}} \frac{1}{\sqrt{n}} \frac{f'}{f}(X_i,\theta)$ to a Brownian motion and thus the existence of δ_o such that :

$$P\left(\underset{k \le n\delta_o}{\text{Sup}} \left|\frac{1}{\sqrt{n}}\sum_{i=1}^{K} \frac{f'}{f}(X_i,\theta_o)\right| > \frac{9A}{2\,I(\theta_o)}\right) \le \frac{\varepsilon}{4} \qquad (T_3)$$

Finally, for k fixed, not greather than

$K(\varepsilon)$, $\sum_{1}^{k} \log \frac{f(X_i,\hat{\theta}_k)}{f(X_i,\theta_o)}$ being a random variable, there exists A_k such that :

$$P\left\{\left|\sum_{i=1}^{k} \log \frac{f(X_i,\hat{\theta}_k)}{f(X_i,\theta_o)}\right| > A_k\right\} < \frac{\varepsilon}{4\,K(\varepsilon)}$$

and for

$$n > N_o = \underset{k=1,\ldots,K(\varepsilon)}{\text{Max}} \left\{ A_k \times \frac{2\ K(\varepsilon)}{A} \right\} \quad \text{we have :}$$

$$P\left(\underset{k=1,\ldots,K(\varepsilon)}{\text{Sup}} \left| \sum_{1}^{k} \log \frac{f\ (X_i,\hat{\theta}_k)}{f\ (X_i,\theta_o)} \ \frac{2\ k}{n} \right| > A \right) \leq \frac{\varepsilon}{4} \qquad (T_4)$$

Now we get the result by combining (T_1), (T_2), (T_3), (T_4).

Corollary 3 : Under assumptions A1 ... A4 and if for each k in $\{1, \ldots, n\}$ equations

$$A : \sum_{i=n-k+1}^{n} \frac{f'}{f}(X_i,\theta) = 0$$

and $$B : \sum_{i=1}^{k} \frac{f'}{f}(X_i,\theta) = 0$$

have an unique solution : $\hat{\theta}^*_k$ (resp $\hat{\theta}_k$), then :

The one dimensional process :

$$\{ Z_n(t) \ , \quad t \in [\ 0,\ 1\ ] \}$$

tends when n tends to infinity, under H_o to

$$\{ B_o^2(t), \quad t \in [\ 0,\ 1\ ] \} \qquad \text{where :}$$

. $B_o(.)$ is the Brownian Brigde

. $Z_n(.)$ is the polygonal line constructed on the points :

$$\{ (0,0), (\frac{k}{n}, \frac{2k}{n} (\frac{n-k}{n}) \left[\sum_{i=1}^{k} \log \frac{f(X_i,\hat{\theta}_k)}{f(X_i,\overset{\times}{\theta}_n)} + \sum_{i=k+1}^{n} \log \frac{f(X_i,\hat{\theta}^*_{n-k})}{f(X_i,\overset{\times}{\theta}_n)} \right]) \ , k=1,\ldots,n$$

Proof : Using arguments of the preceding proofs it is not hard to show that jointly :

$$\{ (Z^*_{n,\theta_o}(t) \quad , \quad Z^{**}_{n,\theta_o}(t)) \quad , \quad t \in [0,1] \}$$

tends to the degenerated process

$$\{((w^2(t) \quad , \quad (w(1)-w(t))^2) \; ; \; t \in [0,1]\}$$

where $Z^{**}_{n,\theta_o}(.)$ is the polygonal line constructed on the points :

$$\{(0,0), \; (\frac{k}{n}, \; 2 \; (\frac{n-k}{n}) \sum_{i=1}^{n-k} \log \frac{f(X_i,\hat{\theta}_{n-k})}{f(X_i,\theta_o)}) \; , \; k=\{1,\ldots,n\}\}$$

To complete the proof it remains to remark that the functional $H : \mathscr{C}[0,1] \to \mathscr{C}[0,1]$

$$f \to H(f)$$

$$H(f)(t) = (1-t) \; f(t) + t(f(1)-f(t)) - t(1-t) \; f(1)$$

is continuous.

3. Likelihood ratio process : convergence under contiguous hypothesis $H^*_{1n}(\tau,\theta_1)$.

It becomes obvious from theorem 1 that H^*_{1n} and H_o are contiguous.

Corollary 3 : Under the assumptions of corollary 2 the process $\{Z^*_{n,\theta_o}(t) \; , \; t \in [0,1]\}$ converges under $H^*_{1n}(\tau,\theta_1)$ to $\{(w(t) + g_{\tau,\theta_1}(t))^2, \quad t \in [0,1]\}$

$$g_{\tau,\theta_1}(t) = \theta_1(1-\tau) \; 1_{t>1-\tau} + \theta_1 t \; 1_{t\leq 1-\tau}$$

Proof : 1) Convergence of the finite dimensional distributions of the Z^*_{n,θ_o} to that of the limit is an easy consequence of the convergence of Z^*_{n,θ_o} under H_o and contiguity of probability measures.

2) Tightness.

Convergence of the Z_n under H_o implies the existence of a modulus of continuity $w(.)$ ($w(\delta) \to 0$ as $\delta \to 0$) such that if $w_n(\delta)$ is the modulus of continuity of the trajectory $Z^*_{n,\theta_o}(.)$ then

$$P_{H_o}\ (\exists \delta \ /\ w_n(\delta) \geq w(\delta) + \varepsilon) \xrightarrow[n\to+\infty]{} 0 \quad , \qquad \forall \varepsilon > 0$$

Contiguity of probability measures implies

$$P_{H^*_{1n}}\ (\exists \delta \ /\ w_n(\delta) \geq w(\delta) + \varepsilon) \xrightarrow[n\to\infty]{} 0 \quad , \qquad \forall \varepsilon > 0$$

and thus implies tightness of the process under $H^*_{1n}(\tau, \delta)$.

BIBLIOGRAPHY.

[1] BAGSHAW M., JOHNSON R.A.:"Sequential procedures for detecting parameter changes in a time-series model" J.A.S.A. (1977), n° 72 p. 593-597

[2] BAHADUR R.R. : "An optimal property of the likelihood ratio statistic" Vth Berkeley Symposium (1966), I, p. 12-26

[3] BHATTACHARYYA G.K., JOHNSON R.A. : "Non parametric tests for shifts at unknown time point", A.M.S. (1968) n° 39, p. 1731-1743

[4] BILLINGSLEY P. : "Convergence of probability measures", Wiley

[5] BIRGE L. : "Vitesses maximales de décroissance des erreurs et tests optimaux associés", Thèse d'Etat (1980), Université Paris VII

[6] BRETAGNOLLE J. : "Statistique de Kolmogorov-Smirnov pour un échantillon non équiréparti", Technical report, Université Paris-Nord

[7] BROWN L.D. : "Non local asymptotic optimality of appropriate likelihood ratio tests", A.M.S. (1971), n° 42, p. 1206-1240

[8] BROWN R.L., DURBIN J., EVANS J.M. : "Techniques for testing the constancy of regression relationships over time", J.R.S.S. (1975), B n° 37; . 149-192

[9] Mc CABE B.P.M., HARRISON M.J. : "Testing the constancy of regression relationships over time using L.S. residuals", App . Stat. (1980), n° 29, p. 142-148

[10] CHERNOFF H. "A measure of asymptotic efficiency for tests based on the sum of observations", A.M.S. (1952), n° 23, p. 493-507

[11] CHERNOFF H., ZACKS S. : "Estimating the current mean of a normal distribution which is subjected to changes in time" A.M.S. (1964),n°35 p. 999-1018

[12] COBB G.W. : "The problem of the Nile : conditional solution to a change-point problem", Biometrika (1978), n° 65, p. 243-251

[13] DELLACHERIE C., MEYER P.A. : "Probabilité et potentiel", Hermann, Paris

[14] DESHAYES J., PICARD D. : "Grandes et moyennes déviations pour les marches aléatoires - applications aux tests de rupture de régression", Astérisque n° 68, ch. IV-V (1979)

[15] DESHAYES J., PICARD D. : "Tests de rupture de régression : comparaison asymptotique", to appear (1981) in Teoriya Veroyatnostei iee Primeneniya (in French).

[16] DVORETZKY A., KIEFER J., WOLFOVITZ J. : "Asymptotic minimax character of the sample distribution", A.M.S. (1956), n° 27, p. 642-669

[17] GALLANT A.R. (1977) : "Testing a non linear regression specification : a non regular case", J.A.S.A. (1977), n° 72, p. 523-530

[18] GARBADE K. : "Two methods for examining the stability of regression coefficients", J.A.S.A. (1977), n° 72, p. 54-63

[19] HAWKINS D.M. : "Testing a sequence of observations for a shift in location", J.A.S.A. (1977), n° 72, p. 180-186

[20] HINKLEY D.V., CHAPMAN P., RUNGER G. : "Change-point problems" invited paper at the 13th European Meeting of Statisticians (1980), Technical report, University of Minnesota

[21] LE CAM L. : "Theorie asymptotique de la décision statistique", Presses de l'Université de Montréal

[22] MOGULSKII A.A. : "Large deviations for trajectories of multidimensional random walks", T.P.A. (1976), n° 21, p. 300-315

[23] Mc NEIL I.B. : "Tests for change of parameter at unknown time and distributions of some related functionals on Brownian motion", A.S. (1974), n° 2, p. 950-962

[24] PAGE R.S. : "A test for a change in a parameter occurring at an unknown point", Biometrika (1955), n° 42, p. 523-526

[25] PETTITT A.N. : "A non parametric approach to the change-point problem" Appl. Stat. (1979), n° 28, p. 126-135

[26] PETTITT A.N. : "A simple cumulative sum type statistic for the change-point problem with zero-one observations", Biometrika (1980), n° 67, p. 79-84

[27] SCHWEDER T. : "Some optimal methods to detect structural shift or outliers in regression", J.A.S.A. (1976), n° 71, p. 491-501

[28] SMITH A.F.M. : "A bayesian approach to inference about a change-point in a sequence of random variables", Biometrika (1975); n° 62, p.407-416

[29] TUSNADY G. : "On asymptotically optimal tests" A.S. (1977), n° 5, p. 385-393

[30] FERNIQUE X. : "Régularité des trajectoires des fonctions aléatoires gaussiennes". Ecole d'Eté de Probabilités de Saint Flour, 1974. Lecture notes in mathematics, n° 480.

[31] BICKEL P.J.- WICHURA M.J. : "Convergence criteria for multiparameter stochastic processes and some applications", A.M.S. n° 42, 1971, p. 1656-1670.

[32] KOMLOS-MAJOR-TUSNADY : "An approximation of Partial sums of independant rv's and the sample DF" : I.Z.Wahrscheinlichtkeitstheorie. Verw. I. Gebiete 32, p. 111-131 (1975), II. Gebiete 34, p. 33-58 (1976).

TESTS DE RUPTURE DE RÉGRESSION : COMPARAISON ASYMPTOTIQUE.

par

J. DESHAYES - D. PICARD

Nous envisageons dans cet article un problème statistique classique où l'emploi de la théorie des grandes déviations permet de simplifier des procédures de choix, d'en suggérer d'autres. Le problème est de déceler l'existence d'une ou plusieurs ruptures dans un modèle de régression indicée de façon chronologique. Nous ne nous préoccupons pas ici des problèmes d'estimation tant des instants de rupture que des paramètres des modèles des différentes phases. Toutefois, les propriétés des différents estimateurs proposés dans la littérature font bien comprendre les difficultés inhérentes au modèle (problèmes de discontinuité de la vraisemblance et problème d'estimation des ruptures proches du début ou de la fin des observations) : citons en particulier HINKLEY [12] et FEDER [9], [10] .

Nous supposerons observer des variables arrivant dans un ordre chronologique préétabli et des ruptures de type quelconque et nous supposerons l'indépendance des variables explicatives et des erreurs. Nous examinerons dans le but de les comparer, essentiellement deux types de tests proposés dans la littérature : les tests du rapport de vraisemblance (voir QUANDT [15], HINKLEY [12], HAWKINS [11]) et les tests sur les marches aléatoires construites sur les résidus récursifs définis par BROWN-DURBIN-EVANS [6]. La difficulté de calculs exacts asymptotiques tant pour le niveau (test du rapport de vraisemblance) que pour la puissance (même en ce qui concerne les tests de résidus récursifs, on ne trouve aucun calcul de puissance dans la littérature) a orienté notre travail vers une comparaison asymptotique de type BAHADUR [2] (ordre de décroissance exponentielle des niveaux et erreurs de deuxième espèce à hypothèses fixes) qu'il convient mieux de considérer comme un outil exploratoire, une première approche, que comme une comparaison stricte.

On montre à travers cette démarche l'optimalité asymptotique du test du rapport de vraisemblance et on est amené à proposer des améliorations sur les tests des résidus récursifs.

Nous proposons de plus une procédure qui introduit une randomisation sur les instants d'observation (les variables observées sont indicées par une variable aléatoire, au lieu d'être indicée par le temps) ramenant ainsi les observations à un échantillon classique d'un problème paramétrique (ce qui en particulier ôte les discontinuités dans la vraisemblance). Les tests qui en découlent ont outre l'intérêt d'être très simples, celui d'être dans certains cas asymptotiquement meilleurs que ceux précédemment étudiés.

Dans la première partie, nous introduisons le modèle et les notations. La seconde partie est consacrée aux tests classiques fondés sur les résidus récursifs, les rapports de vraisemblance et traite la procédure de randomisation. La troisième compare les différents tests. Enfin, la quatrième partie est consacrée aux techniques de grandes déviations utilisées dans les calculs d'erreurs des différents tests, et aux problèmes concrets qui se sont posés à ce propos.

I. MODELE et NOTATIONS.

Soit le modèle

$$y_k = x_k' \beta_k + u_k \ , \quad k = 1,\ldots,n$$

où x_k désigne le vecteur de dimension s sur lequel on régresse, il est considéré comme une constante (ou bien aléatoire mais indépendant des erreurs) ; x_k' est son transposé.

β_k désigne le vecteur de dimension s des paramètres de la régression au temps k.

u_k est l'erreur ; les u_k sont supposées indépendantes et de loi gaussienne $\mathcal{N}(0,\sigma_k^2)$.

L'hypothèse à tester est qu'il n'y a pas de rupture : $H_o = \{\beta_1 = \beta_2 = \ldots = \beta_n$ et $\sigma_1 = \sigma_2 = \ldots = \sigma_n\}$. Les résidus récursifs (ou innovations en filtrage) sont les variables :

$$w_k = \frac{y_k - x_k' B_{k-1}}{\sqrt{1 + x_k' (X_{k-1}' X_{k-1})^{-1} x_k}} \quad \text{pour} \quad k = s+1,\ldots,n$$

avec

$$X_k' = (x_1,\ldots,x_k)$$

$$Y_k' = (y_1,\ldots,y_k)$$

$$B_k = (X_k' X_k)^{-1} X_k' Y_k$$

qui sont obtenues à partir des résidus habituels par orthogonalisation. Si on y ajoute les s premières observations, ces résidus récursifs sont exhaustifs.

Sous l'hypothèse H_o, ils sont normaux $\mathcal{N}(0,\sigma_1^2)$ et indépendants.

Sous une contre-hypothèse où il y a changement de paramètre β, le paramètre σ restant le même, ces résidus gardent la même structure de covariance mais ils sont biaisés après la rupture. Sous une contre-hypothèse où il y a changement de paramètre σ, le paramètre β restant le même, les résidus restent sans biais mais leur covariance est modifiée.

On voit donc que tout problème de rupture dans un modèle de ce type se ramène à un problème de rupture sur l'aspect : n-échantillon gaussien des variables w_k.

Bien que la plupart des tests étudiés soient utilisés pour déceler des ruptures de type général, nous nous limiterons pour les calculs de puissance, à des contre-hypothèses H_1 du type :

$$\begin{cases} \exists\, k_o = 1,2,\ldots,n \text{ tel que } \beta_1 = \ldots = \beta_{k_o} \neq \beta_{k_o+1} = \ldots = \beta_n \\ \qquad\qquad\qquad\qquad \sigma_1 = \sigma_2 = \ldots = \sigma_n. \end{cases}$$

Nous ferons les calculs de vitesses exponentielles des erreurs dans le cas d'une rupture de moyenne dans un n-échantillon gaussien. Toutefois, pour les erreurs de deuxième espèce, les biais dépendant des régresseurs : nous indiquerons alors ce qui se passe dans des cas particuliers importants (les x_k pouvant être considérés comme des variables aléatoires indépendantes identiquement distribuées, ou des fonctions polynômiales du temps).

De point de vue asymptotique, nous supposerons que l'observation $(y_1,\ldots,y_n)$ provient en fait d'un processus Y_t défini pour t compris entre 0 et 1, que l'on observe aux instants $(\frac{1}{n}, \frac{2}{n}, \ldots, 1)$. Ce point de vue correspond à une observation de plus en plus fine d'un phénomène, comme on peut le faire en exploration automatique de photos par exemple. Quand n est fixé, on confondra par abus de langage : y_k et $Y_{\frac{k}{n}}$.

Sous l'alternative, on observe une rupture de moyenne dans le processus à un instant t_o inconnu compris strictement entre 0 et 1. Pour une observation de taille n, on a :

. pour $k = 1,2,\ldots, [nt_o]$, $Y_{\frac{k}{n}}$ suit une loi $\mathcal{N}(\beta^*,\sigma^2)$

. pour $k = [nt_o]+1,\ldots,n$, $Y_{\frac{k}{n}}$ suit une loi $\mathcal{N}(\beta^{**},\sigma^2)$.

De sorte que l'on réfèrera un point de H_o par le couple (β,σ) et un point de H_1 par le quadruplet $(\beta^*,\beta^{**},\sigma,t_o)$.

Il reste à définir le sens dans lequel vont s'effectuer les comparaisons entre les différents tests. Les hypothèses H_o et H_1 (défnies comme précédemment) sont considérées comme fixes (on ne les fait pas se rapprocher) ; les erreurs de première et de deuxième espèce décroissent exponentiellement quand le test est bon, les hypothèses suffisamment séparées et le seuil bien choisi (cf. BAHADUR [8], TUSNADY [32]).

De façon plus précise, si $\{D_n^i\}$ est la suite de régions de rejet du test $\mathcal{T}_i$ telle que

$$\lim_{n\to\infty} \frac{1}{n} \log P_{H_o}(D_n^i) \to -\alpha \qquad i = 1,2$$

en désignant par $\bar{D}_n^i$ le complémentaire de la région de rejet, on a :

$$\lim_{n\to\infty} \frac{1}{n} \log P_{H_1}(\bar{D}_n^i) \to -\gamma(\mathcal{T}_i).$$

On dira qu'au niveau de décroissance α, $\mathcal{T}_1$ est meilleur que $\mathcal{T}_2$ si : $\gamma(\mathcal{T}_1) > \gamma(\mathcal{T}_2)$.

Toutefois, il arrivera que les expressions des erreurs de deuxième espèce soient trop compliquées pour autoriser ce genre de comparaison ; à niveau de décroissance α fixé, on comparera alors les intervalles respectifs de décroissance exponentielle de l'erreur de deuxième espèce (fonction des paramètres de l'hypothèse $H_1 : t_o$ et $\Delta = \frac{|\beta^* - \beta^{**}|}{\sigma}$). On notera alors qu'apparaissent des expressions du type : $t_o(1-t_o)\,\Delta^2 = \ell_{t_o}^2$ (ou semblables), ce qui semblerait indiquer que ℓ serait le vrai paramètre du problème.

II. TESTS CLASSIQUES CONSTRUITS A L'AIDE DES RESIDUS RECURSIFS.

Dans l'espace $\mathscr{C}[0,1]$ des fonctions continues sur $[0,1]$, muni de la norme uniforme, on notera $s_{\xi,n}(.)$ la ligne polygonale joignant les points $(0,0)$ et $(\frac{k}{n}, \frac{\sum_{i=1}^{k} \xi_i}{n})_{k=1,\ldots,n}$, où $(\xi_1,\ldots,\xi_n)$ est une famille de variables aléatoires.

II.1. Tests classiques sur les sommes cumulées.

II.1.a) Définition des tests :

Dans le cas σ connu, les tests classiques consistent à rejeter l'hypothèse H_o lorsqu'une marche aléatoire $s_{\xi,n}(.)$ sort du domaine limité par 2 fonctions continues $g_1(.)$ négative et $g_2(.)$ positive :

$$D = \{\exists\, t \in [0,1],\ s_{\xi,n}(t) < g_1(t) \ \text{ou}\ s_{\xi,n}(t) > g_2(t)\}$$

avec $\xi_i = f(w_i)$.

Les fonctions f employées dans [6] sont $f(w) = \frac{w}{\sigma}$ pour le test de la somme cumulée des résidus (SCR) et $f(w) = \frac{w^2}{\sigma^2}$ pour le test de la somme cumulée des carrés des résidus (SCCR) associées respectivement aux frontières : $g_2(t) = -g_1(t) = at + \frac{a}{2}$ pour le test (SCR) et $g_1(t) \equiv 0$, $g_2(t) = t+c$ pour le test (SCCR).

Dans le cas où σ est inconnu, on le remplace par l'estimateur des moindres carrés : $\hat{\sigma}^2 = s_{w^2,n}(1)$.

II.1.b) Choix des frontières.

Dans [6], les frontières choisies l'ont été parce qu'elles sont simples et permettent le calcul du niveau des tests : de façon approchée pour tout n (test SCCR) ou asymptotiquement (test SCR) en approchant la marche aléatoire $\frac{\sqrt{n}}{\sigma} s_{w,n}(.)$ par un mouvement brownien. Malheureusement, cette approximation entraîne une erreur sur le niveau qui est pour des échantillons de taille moyenne du même ordre que le niveau lui-même : l'erreur absolue sur la probabilité calculée étant $O(\frac{1}{\sqrt{n}})$ pour un nombre n d'observations égal à 200, on s'aperçoit par simulation que le niveau effectif est la moitié du niveau supposé. Pour des nombres d'observations n petits, un calcul direct de probabilité de gaussienne dans $\mathbb{R}^n$ est préférable.

Parmi les tests associés à la même fonction f des résidus, on a le choix entre une infinité de frontières g_i possibles. Aucune n'est uniformément meilleure que les autres.

L'abord du problème par les méthodes des grandes déviations a l'avantage de fournir des frontières g (asymptotiquement) optimales pour chaque type de test à cause de la remarque suivante :

Remarque. Supposons qu'il existe une famille de fonctions $(H_t)_{t\in]0,1]}$ telles que :

$$\begin{cases} H_t(a) \leq 0, \ \forall\ a \in I \subset \mathbb{R}, \ \forall\ t \in]0,1] \\ H_t(.) \text{ est décroissante sur } I \text{ (non identiquement nulle)}, \ \forall\ t \in]0,1] \\ \lim\limits_{n\to\infty} \frac{1}{n} \log P(s_{\xi,n}(t) > a) = H_t(a) \\ \lim\limits_{n\to\infty} \frac{1}{n} \log P(\exists\, t \in [0,1],\ s_{\xi,n}(t) > g(t)) = \sup\limits_{t\in[0,1]} H_t[g(t)]. \end{cases}$$

Alors, parmi toutes les frontières g donnant un niveau exponentiel α, il y en a une qui définit une région de rejet contenant toutes les autres : c'est la fonction g^* définie pour tout t compris entre 0 et 1 par

$$g^*(t) = \inf\{y, y \in I,\ H_t(y) = -\alpha\}\ .$$

Il est intéressant de remarquer que ces frontières g^* optimales répondent asymptotiquement à une demande intuitive souvent formulée : que la probabilité d'avoir franchi la frontière à un instant $\frac{k}{n}$ est la même pour tout $k = 1,2,\ldots,n$ (sous H_0).

Pour le test basé sur les sommes cumulées de résidus (SCR), les fonctions g^* optimales sont les paraboles : $g_2^*(t) = c\sqrt{t} = -g_1^*(t)$.

Pour le test basé sur les sommes de carrés de résidus (SCCR), les fonctions g_1^* et g_2^* optimales sont :

- dans le cas σ connu : solutions de l'équation implicite pour un niveau exponentiel α fixé

$$\frac{g_1^*(t)-2\alpha}{t} = \log[1 + \frac{g_1^*(t)}{t}]$$

- dans le cas σ inconnu solutions de l'équation :

$$\begin{cases} g_2^*(t) = g_1^*(1-t) \\ t \log[1 + \frac{g_1^*(t)}{t}] + (1-t)\log[1 - \frac{g_1^*(t)}{1-t}] = -2\alpha. \end{cases}$$

II.2. Test du rapport de vraisemblance : optimalité.

Pour simplifier l'exposition, plaçons-nous dans le cas où σ est connu. Soit $L(\beta)$ la vraisemblance en un point β de H_o, $L(\beta^*,\beta^{**},t_o)$ celle calculée en un point (β^*,β^{**},t_o) de H_1.

Le test Φ_n du rapport de vraisemblance est défini par la région de rejet de H_o :

$$\mathcal{D}_n = \{\frac{1}{n} \log \frac{\sup_{H_1} L(\beta^*,\beta^{**},t_o)}{\sup_{H_o} L(\beta)} \geq c\} = \{\exists\, k = 2,\dots,n-1 \text{ tel que } |\frac{\bar{y}_{1,k} - \bar{y}_{k+1,n}}{\sigma}| \geq \sqrt{\frac{2c}{\frac{k}{n}(1-\frac{k}{n})}}\}$$

où $\bar{y}_{k,\ell} = \frac{1}{\ell-k-1} \sum_{i=k+1}^{\ell} y_i$.

PROPOSITION.- Pour tout $\alpha > 0$, le test du rapport de vraisemblance (à σ positif connu) est uniformément exponentiellement optimal parmi tous les tests de niveau exponentiel au moins α .

i.e. Parmi toutes les suites de tests $\{\varphi_n\}$ satisfaisant

$$\sup_{\beta \in H_o} \limsup_{n\to\infty} \frac{1}{n} \log E_\beta \varphi_n \leq -\alpha$$

Φ_n est tel que : $\forall\, (\beta^*,\beta^{**},t_o) \in H_1$:

$$\liminf_{n\to\infty} \frac{1}{n} \log(1 - E_{(\beta^*,\beta^{**},t_o)} \varphi_n) \geq \limsup_{n\to\infty} \frac{1}{n} \log\left(1 - E_{(\beta^*,\beta^{**},t_o)} \Phi_n\right).$$

Démonstration.- La région de rejet $\mathscr{D}_n$ est l'union de n-1 régions $\mathscr{D}_n(\frac{k}{n})$ qui constituent, chacune, la région de rejet du test de Neyman-Pearson de l'hypothèse simple de H_o : $\beta = \frac{k}{n}\beta^* + (1-\frac{k}{n})\beta^{**}$ contre l'hypothèse simple de H_1 : $(\beta^*,\beta^{**},\frac{k}{n})$. A $\frac{k}{n} = t$ fixé, le test $\mathscr{D}_n(t)$ est exponentiellement optimal comme test de $\{\beta\}$ contre $\{(\beta^*,\beta^{**},t_o)\}$ en raison de son optimalité à distance finie (parmi les tests sans biais) et de la continuité des fonctions e_1 et e_2 :

$$e_1 : c \rightarrow e_1(c) = \lim_{n\to\infty} \frac{1}{n} \log P_\beta \ (\mathscr{D}_n(t)).$$

$$e_2 : c \rightarrow e_2(c) = \lim_{n\to\infty} \frac{1}{n} \log P_{(\beta^*,\beta^{**},t)}(\overline{\mathscr{D}_n(t)}).$$

(Il suffit de calculer les expressions).

Il est aussi exponentiellement optimal, comme test de H_o (tout entier) contre (β^*,β^{**},t). (En effet, le point $\beta = t\beta^* + (1-t)\beta^{**}$ est le point le plus défavorable de H_o contre (β^*,β^{**},t) : celui, qui, à vitesse de décroissance fixée pour l'erreur de 2ème espèce, donne le niveau le plus grand).

De plus, les (n-1) régions $\mathscr{D}_n(\frac{k}{n})$ ayant toutes le même niveau asymptotique, le test global a même niveau et une erreur de 2ème espèce meilleure que chacun d'eux et donc égale au meilleur d'entre eux :

$$\lim_{n\to\infty} - \frac{1}{n} \log P_\beta \ (\mathscr{D}_n) = c$$

$$\lim_{n\to\infty} - \frac{1}{n} \log P_{(\beta^*,\beta^{**},t_o)}(\overline{\mathscr{D}}_n) = \begin{cases} \dfrac{(\ell_{t_o} - \sqrt{2c})^2}{2} & \text{si } \ell_{t_o}^2 \geq 2c \\ 0 & \text{sinon} \end{cases}$$

(où $\ell_{t_o}^2$ désigne : $t_o(1-t_o)\left(\frac{\beta^*-\beta^{**}}{\sigma}\right)^2$).

Dans le cas σ inconnu, le test de vraisemblance a, alors, pour région de rejet

$$\mathscr{D}_n = \{\exists\, k = 2,\dots,n-1,\ \frac{|\bar{y}_{1,k} - \bar{y}_{k+1}|}{\sqrt{\frac{SS_{1,k}+SS_{k+1,n}}{n}}} \geq \frac{c}{\sqrt{\frac{k}{n}(1-\frac{k}{n})}} \}$$

où $SS_{k,\ell} = \sum_{j=k}^{\ell} (y_j - \bar{y}_{k,\ell})^2$.

L'optimalité pour n fixé, k connu, du test de Student parmi les tests invariants (pour le groupe des homothéties translations), permet de déduire la proposition.

PROPOSITION.- Pour tout $\alpha > 0$, le test du rapport de vraisemblance est uniformément, exponentiellement optimal parmi les tests invariants, de niveau exponentiel au moins α .

Remarque. Ce qui change dans le cas d'une régression générale :

Dans le modèle $y_k = x_k'\beta_k + u_k$, $k = 1,\ldots,n$, si on suppose que la première composante contient l'espérance : $x_{k1} = 1$ et que les vecteurs $(x_{k2},\ldots,x_{ks})$, $k = 1,\ldots,n$ sont aléatoires, indépendants (et indépendants des u_k) de même loi : centrés et de matrice de covariance Γ, le calcul des biais des résidus est le même que pour une simple rupture de moyenne : la loi des grands nombres montre qu'il suffit de remplacer $(\beta^{**}-\beta^*)$ par sa première composante en ce qui concerne $s_{w,n}(.)$ et $(\beta^{**}-\beta^*)^2$ par $(\beta^{**}-\beta^*)' \begin{pmatrix} 1 & 0 \\ 0 & \Gamma \end{pmatrix} (\beta^{**}-\beta^*)$ en ce qui concerne $s_{w^2,n}(.)$. L'optimalité du test du rapport de vraisemblance demeure parmi les tests invariants par homothétie-translation.

II.3. Etude d'un test avec randomisation sur les instants d'observations dans le cas d'une rupture de moyenne.

La lecture des études consacrées au comportement de la vraisemblance dans ces problèmes de ruptures, met en valeur certaines difficultés sous-jacentes (discontinuités de la vraisemblance, problèmes de bords). On va essayer de contourner ces difficultés en changeant la procédure d'observation : au lieu d'observer le processus Y_t à des instants $\frac{k}{n}$ fixes, on va tirer d'abord et indépendamment, au sort les instants d'observations par exemple avec un n-échantillon $(U_1,\ldots,U_n)$ d'une loi π sur $[0,1]$, de sorte que, si π est une loi sans masse, on obtient un n-échantillon $(Z_1,Z_2,\ldots,Z_n)$, d'une loi normale $N(\beta,\sigma^2)$ sous l'hypothèse H_o , d'une loi mélange : $N(\beta^*,\sigma^2)$ (proportion π_o) et $N(\beta^{**},\sigma^2)$ $(1-\pi_o)$ $(\pi_o = \pi(U < t_o))$ sous l'hypothèse H_1. On est alors ramené à un problème paramétrique classique avec un échantillonnage et on peut employer les méthodes habituelles.

L'évaluation du 2^e moment (sous H_o : σ^2 , sous H_1 : $\sigma^2(1+\ell^2_{\pi_o})$ où $\ell^2_{\pi_o} = \pi_o(1-\pi_o)\frac{|\beta^*-\beta^{**}|^2}{\sigma^2}$) fait apparaître $\ell^2_{\pi_o}$ comme paramètre du problème, et conduit naturellement à un test consistant en une comparaison d'un estimateur de la

variance de l'échantillon avec σ^2 (connu ou estimé par une autre méthode).

Dans le cas σ connu, la région de rejet est donc

$$D = \{ \frac{1}{(n-1)\sigma^2} \sum_{i=1}^{n} (Z_i - \bar{Z})^2 \geq 1+c \} .$$

Le calcul simple des erreurs exponentielles (IV.1) donne des indications sur la loi π à choisir :

$$\forall\, c > 0 \qquad \lim_{n\to\infty} -\frac{1}{n} \log P_{H_o}(D) = \frac{c - \log(1+c)}{2}$$

$$\forall\, c < \ell^2_{\pi_o} \qquad \lim_{n\to\infty} -\frac{1}{n} \log P_{H_1}(\bar{D}) = \mathscr{H}_{\pi_o}(c) .$$

L'expression de $\mathscr{H}_{\pi_o}(c)$ (cf. IV.1) met en évidence des inégalités du type :

$$\mathscr{H}_{\pi_o}(c) = \mathscr{H}_{1-\pi_o}(c)$$

$$\forall\, \pi, \pi_o \,/\, \pi_o \leq \pi \leq \tfrac{1}{2} \quad \mathscr{H}_{\pi_o}(c) \leq \mathscr{H}_{\pi}(c) \leq \mathscr{H}_{\frac{1}{2}}(c) .$$

On aura donc intérêt à choisir la loi π, la plus simple possible, mais dans la mesure où l'on dispose d'information a priori sur t_o, on aura intérêt à introduire une procédure qui "tire autant d'observations avant t_o qu'après".

Remarque. Dans le cas σ inconnu, on conserve le même principe de test en remplaçant σ^2 par exemple par une estimation faite à l'aide d'un deuxième échantillon de taille n : $V_1,\ldots,V_n$ réalisé suivant une procédure analogue à celui de $(Z_1,\ldots,Z_n)$ des instants d'observations étant tirés suivant une loi $\pi'^{\otimes n}_n$ de sorte que la variance σ^2 soit estimée sans biais pour n assez grand. (Un choix théorique possible pour π'_n est une loi uniforme sur $[0,a_n]$, a_n décroissant vers 0). Pratiquement on aura plutôt intérêt à construire π'_n en fonction de l'information a priori dont on dispose.

En réalité tant au niveau théorique que pratique, on aura intérêt à procéder un peu différemment : n étant fixé, on suppose disposer de $N(n)$ instant d'observations possibles disposés sur $[0,1]$ (avec $\frac{N(n)}{n} \xrightarrow[n\to\infty]{} \infty$). Soient $u_1,u_2,\ldots,u_N$ ces instants possibles et $Y_{u_1}, Y_{u_2},\ldots, Y_{u_N}$ leurs observations associées. On effectuera un tirage uniforme et sans remise de n instants parmi les $N(n)$. L'arbitraire de la loi π se retrouve dans le choix des instants $u_1,\ldots,u_N$. Les deux procédures

sont asymptotiquement équivalentes pour le problème qui nous intéresse.

III. COMPARAISON DES DIFFERENTS TYPES DE TESTS.

III.1. Comparaison du test SCCR avec le test à procédure randomisée (σ connu).

La comparaison de ces deux tests se justifie à de nombreux points de vue.

En effet, ils sont tous les deux du type "analyse de variance", l'un inspecte à chaque pas :

$$\sum_{i=2}^{k} \omega_i^2 = \sum_{i=2}^{k} (y_i - \bar{y}(1,k))^2 \quad k = 2,\dots,n \quad (\bar{y}(1,k) = \frac{1}{k}\sum_{i=1}^{k} y_i)$$

l'autre n'inspecte que la dernière somme $\sum_{i=1}^{n} (z_i - \bar{z})^2$. On pourrait donc à première vue penser que le premier est de ce fait plus intéressant. En fait,

sous H_0 , les variables y_i et z_i ont même loi mais les expressions des erreurs de 2^e espèce asymptotiques sont assez semblables.

Il est en particulier facile de montrer que pour tout c positif, il existe une valeur $\pi_0(c)$ telle que $\pi_0(c) \leq \inf(t_0, 1-t_0)$ et pour toute loi π vérifiant : $\pi_0(c) \leq \pi(U < t_0) \leq 1 - \pi_0(c)$, le test à procédure randomisée est exponentiellement meilleur que le test SCCR (voir les expressions obtenues au § IV.1).

Il est intéressant de noter que la randomisation améliore le test même si la loi π est uniforme et qu'elle pourra l'améliorer d'autant plus si on dispose d'information a priori sur la localisation de l'instant de rupture permettant de "dilater" ou "contracter" l'intervalle $[0,t_0]$ à l'aide de la loi π pour le rapprocher de $[0,\frac{1}{2}]$.

III.2. Comparaison globale des tests sur leurs zones de décroissance exponentielle (σ connu).

Les expressions obtenues (voir § IV) pour les erreurs de deuxième espèce des différents tests sont assez compliquées et ne permettent pas de faire dans tous les cas des comparaisons strictes. En particulier, nous n'envisagerons pas le cas σ inconnu, conjecturant que les qualités des tests ne se modifient pas trop par l'estimation de σ . On prendra aussi en compte l'étendue des zones de décroissances exponentielles des erreurs de deuxième espèce : en effet, savoir qu'un test ne décroît pas

exponentiellement sur une certaine plage est a priori plus informatif sur ses qualités qu'une comparaison exacte d'ordres de décroissance exponentielle.

Le test du maximum de vraisemblance a évidemment, grâce à son optimalité, une zone de décroissance exponentielle plus étendue. La vitesse exponentielle du niveau étant fixée égale à $\alpha = \frac{c-\log(1+c)}{2}$, les différents tests ont une décroissance exponentielle de l'erreur de 2^e espèce si :

$\Delta > \frac{\sqrt{\alpha}}{\mu(t_o)}$ ou $\Delta > \frac{\sqrt{\alpha}}{\frac{2}{3}t_o \log \frac{1}{t_o}}$ pour le test basé sur les sommes de résidus

$\Delta > \sqrt{\frac{c}{t_o(1-t_o)}}$ pour le test basé sur les sommes de carrés

$\Delta > \sqrt{\frac{2\alpha}{t_o(1-t_o)}}$ pour le test du maximum de vraisemblance

$\Delta > \sqrt{\frac{c}{\pi_o(1-\pi_o)}}$ pour le test avec randomisation.

Comme on peut le constater sur les expressions ci-dessus, le test (SCCR) est meilleur "aux bords" que le test (SCR) et est moins bon pour les valeurs moyennes de t_o.

Le test de procédure de randomisation a une région moins étendue que le test du maximum de vraisemblance lorsque la loi π est uniforme mais qui peut devenir plus étendue si l'information a priori conduit à prendre une loi π vérifiant :

$$\pi_o(1-\pi_o) > t_o(1-t_o)\,\frac{c}{c-\log(1+c)}$$

ce qui n'est d'ailleurs possible que pour des valeurs de t_o voisines de 0 ou 1.

IV. ANNEXE DES DEMONSTRATIONS. THEOREME DES GRANDES DEVIATIONS.

IV.1. Grands écarts dans $\mathbb{R}$.

On reprend ici une démonstration des inégalités de CHERNOFF [7] adaptées aux problèmes que l'on se pose ici (voir aussi DACUNHA-CASTELLE [8]). Soit Z_n une suite de variables aléatoires réelles de lois F_n, et a un nombre fixé positif. Soit $\varphi_n(.)$ la transformée de Laplace de Z_n et h_n la conjuguée de YOUNG-ORLICZ [13] de $\frac{1}{n}\log\varphi_n$:

$$\varphi_n(t) = E_{F_n} \exp t Z_n$$

$$h_n(x) = \sup_{t \in \mathbb{R}} xt - \frac{1}{n} \log \varphi_n(t).$$

IV.1.a) Théorème.

Sous les hypothèses i → iv) suivantes, $\lim_{n\to\infty} \frac{1}{n} \log P(Z_n > na) = -h(a)$

i) $EZ_n < 0$;

ii) $\frac{1}{n} \log \varphi_n(t)$ converge vers une fonction $\psi(t)$ en tout point intérieur au domaine de définition de ψ .

iii) 0 est un point intérieur du domaine de définition de ψ .

iv) a est un point intérieur du domaine de définition de h (h est la conjuguée de Young-Orlicz de ψ).

Démonstration.- Il est démontré dans BARTFAI [3] que ces conditions impliquent : d'une part la convergence de h_n vers h sur le domaine de définition de h, d'autre part si t_n est la suite définie par : $h_n(a) = a t_n - \frac{1}{n} \log \varphi_n(t)$ alors il existe t appartenant au domaine de définition de ψ tel que : t_n converge vers t et $h(a) = at - \psi(t)$.

Démontrons maintenant l'égalité du théorème. On a d'abord

$$P_{F_n}(Z_n > na) \leq E \exp t(Z_n - na), \quad \forall \ t > 0$$

d'où $$\overline{\lim_{n\to\infty}} \frac{1}{n} \log P(Z_n > na) \leq -h(a).$$

D'autre part, si b est un réel strictement plus grand que a et F_{t_n} la loi de probabilité qui admet la densité $\frac{e^{t_n x}}{\varphi(t_n)}$ par rapport à F_n , t_n étant ici défini par : $h_n(b) = bt_n - \frac{1}{n} \log \varphi_n(t_n)$, on a :

$$P_{F_n}(Z_n > na) = \varphi_n(t_n) \, E_{F_{t_n}} e^{-t_n Z_n} 1_{Z_n - na > 0}$$

d'où

$$\frac{1}{n} \log P_{F_n}(Z_n > na) \geq -bt_n + \frac{1}{n} \log \varphi_n(t_n) + \frac{1}{n} \log \left\{ E_{F_{t_n}} e^{-nt_n \frac{Z_n - nb}{n}} \times 1_{-\varepsilon < \frac{Z_n - nb}{n} < \varepsilon} \right\}$$

avec $\varepsilon = b - a$.

On a en effet convergence dans L^2 de $\frac{Z_n - nb}{n}$ vers 0 sous la loi F_{t_n}.

On obtient donc :

$$\frac{1}{n} \log P_{F_n}(Z_n > na) \geq - \inf_{b>a} h(b).$$

Le résultat se déduit de la convexité de h et de l'inégalité $b > EZ_n$ pour tout n.

IV.1.b) Applications : niveau et erreur de 2^e espèce exponentiels du test avec randomisation lorsque σ est connu.

$$\lim_{n\to\infty} \frac{1}{n} \log P_{H_0}\left\{ \frac{\sum_{i=1}^{n} (X_i-\bar{X})^2}{(n-1)\,\sigma^2} > 1+c \right\} = -\mathscr{H}_0(c), \quad \forall\, c>0$$

avec $\mathscr{H}_0(c) = \dfrac{c-\log(1+c)}{2}$

$$\lim_{n\to\infty} \frac{1}{n} \log P_{H_1}\left\{ \frac{\sum_{i=1}^{n} (X_i-\bar{X})^2}{(n-1)\,\sigma^2} < 1+c \right\} = -\mathscr{H}_{\pi_0}(c), \quad \forall\, c < \ell^2_{\pi_0}$$

avec $\mathscr{H}_{\pi_0}(c) = \dfrac{c+1}{2}\,[1-\Lambda(\alpha)] - \dfrac{\alpha^2+(1-\alpha)^2}{4}\left[\dfrac{1}{\Lambda(\alpha)} - 1\right] \Delta^2 + \dfrac{1}{2} \log \Lambda(\alpha) - \log \dfrac{(1-\pi_0)\pi_0}{(1-\alpha)\alpha}$,

où $\Delta = \dfrac{|\beta^* - \beta^{**}|}{\sigma}$, $\Lambda(\alpha) = \dfrac{1+\sqrt{1+4(c+1)\ell^2_\alpha}}{2(c+1)}$, et $\alpha = \alpha(c)$ est la solution de l'équation

$$(*) \qquad \log \frac{\pi_0}{\alpha} \frac{1-\alpha}{1-\pi_0} + (1-2\alpha)\left[\frac{1}{\Lambda(\alpha)} - 1\right] \frac{\Delta^2}{2} = 0 .$$

IV.2. Grands écarts dans $\mathbb{R}^k$ (coordonnées indépendantes).

IV.2.a) Enoncé du théorème.

Si $Z_n = (Z_n^1, \ldots, Z_n^k)$ chaque suite $(Z_n^i)_{n\in\mathbb{N}}$ vérifiant les conditions i $\rightarrow$ iv) du paragraphe précédent et (Z_n^i) étant indépendante de (Z_n^j), pour $i \neq j$.

Alors on a pour tout borélien A de $\mathbb{R}^k$;

$$H(\bar{A}) \leq \underline{\lim_{n\to\infty}} - \frac{1}{n} \log P\left(\frac{Z_n}{n} \in A\right) \leq \overline{\lim_{n\to\infty}} - \frac{1}{n} \log P\left(\frac{Z_n}{n} \in A\right) \leq H(\overset{\circ}{A})$$

où $H(A) = \inf_{x\in A} h(x)$

$$h(x) = \sum_{i=1}^{k} h_i(x_i) \quad \text{pour} \quad x = (x_1, \ldots, x_k).$$

La démonstration du passage de $\mathbb{R}$ à $\mathbb{R}^k$ pour les inégalités de Chernoff que l'on trouve, par exemple, dans [26] se transpose point par point à ce cas simple où les coordonnées sont indépendantes.

IV.2.b) Applications : niveau et erreur de 2^{e} espèce exponentielle du test avec randomisation (σ inconnu).

$$\lim_{n\to\infty} -\frac{1}{n}\log P_{H_0}\left\{\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{\sum_{i=1}^{n}(V_i-\bar{V})^2} > 1+c\right\} = \log\frac{1+\frac{c}{2}}{\sqrt{c}}, \quad \forall\ c>0$$

$$\lim_{n\to\infty} -\frac{1}{n}\log P_{H_1}\left\{\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{\sum_{i=1}^{n}(V_i-\bar{V})^2} > 1+c\right\} = \mathscr{H}_1(c), \quad \forall\ c < \ell_{\pi_0}^2$$

avec

$$\mathscr{H}_1(c) = \frac{1}{2}\log\Lambda(\alpha) - \frac{\alpha^2\Delta^2}{2}\left[\frac{1}{\Lambda(\alpha)} - 1\right] + \frac{1}{2}\log\left[1+(c+1)(1-\Lambda(\alpha))\right] - \log\frac{1-\pi_0}{1-\alpha}$$

$$\Lambda(\alpha) = \frac{c+2-\tau(\alpha,c) - \sqrt{(c+2)^2+6\tau(\alpha,c)(c+2) + [\tau(\alpha,c)]^2}}{4(c+1)}$$

avec $\tau(\alpha,c) = (c+1)\ell_\alpha^2$ et α solution de (*).

Les théorèmes de grandes déviations sur $\mathbb{R}^k$ permettent d'étendre les résultats aux marches aléatoires : dans le cas stationnaire, on se reportera à [4] et [14] ; dans le cas non stationnaire, on utilisera les résultats de VENTSEL [18] énoncés aussi dans [1].

IV.3. Grands écarts pour les marches aléatoires dans le cas stationnaire.

IV.3.a) Enoncé du théorème [14].

Si les vecteurs aléatoires $\xi_1, \ldots, \xi_n \ldots$ centrés ont une transformée de Laplace h_ξ définie sur un voisinage de 0 dans $\mathbb{R}^m$; la marche aléatoire $s_{\xi,n}(.)$ est telle que :

$$\limsup_{n\to\infty}\frac{1}{n}\log P(s_{\xi,n}(.)\in G) = \liminf_{n\to\infty}\frac{1}{n}\log P(s_{\xi,n}(.)\in G) = -W_\xi(G)$$

pour les boréliens G de $\mathscr{C}_m$ vérifiant $W_\xi(\mathrm{ad}(G)) = W_\xi(\mathrm{int}(G))$ où $\mathrm{int}(G)$ désigne l'intérieur de G et $\mathrm{ad}(G)$ son adhérence relativement à la topologie trace de Skorokhod sur $\mathscr{C}_m[0,1]$.

IV.3.b) Application à un franchissement de barrière (en dimension 1).

Pour une suite g_n de fonctions continues sur $[0,1]$, qui converge vers une fonction g continue, strictement positive sur $[0,1]$ et telle qu'il existe τ compris entre 0 et 1 tel que

$$\text{ess.inf } \xi < \frac{g(\tau)}{\tau} < \text{ess sup } \xi ,$$

on définit le borélien $G_{g_n} = \{f \in \mathscr{C}_{[0,1]} , \exists\, t \in [0,1] \text{ tel que } f(t) > g_n(t)\}$, alors

$$\lim_{n\to\infty} -\frac{1}{n} \log P\{s_{\xi,n}(.) \in G_{g_n}\} = \inf_{t\in[0,1]} t\, h[\frac{g(t)}{t}] .$$

De façon évidente, le résultat reste encore vrai lorsque $g(0) = 0$ mais $g'(0) > \text{ess sup } \xi$ et si $\lim_{t\to 0} \frac{g_n(t)}{t} = \lim_{t\to 0} \frac{g(t)}{t}$.

Les niveaux des tests (SCR) et (SCCR) sont, pour σ connu, respectivement donnés par :

$$\lim_{n\to\infty} -\frac{1}{n} \log P_{H_o}\{\exists\, t \in [0,1],\ \frac{s_{w,n}(t)}{\sigma} > at + \frac{a}{2}\} = a^2$$

$$\lim_{n\to\infty} -\frac{1}{n} \log P_{H_o}\{\exists\, t \in [0,1],\ \frac{s_{w^2,n}(t)}{\sigma^2} > t+c\} = \frac{c - \log(1+c)}{2} .$$

Une extension du problème de franchissement de barrières consiste pour des variables aléatoires ξ telles que $E\xi = a > 0$ et ess sup $\xi = +\infty$ à considérer l'ensemble $G_g = \{f \in \mathscr{C}[0,1],\ \exists\, t \in [0,1] \text{ tel que } \frac{f(t)}{f(1)} > g(t)\}$. Alors

$$\lim_{n\to\infty} -\frac{1}{n} \log P\{s_{\xi,n}(.) \in G_g\} = \inf_{\lambda>0} \inf_{t\in[0,1]} \{t.h[\frac{\lambda g(t)}{t}]+(1-t)\, h[\frac{\lambda(1-g(t))}{1-t}]\}$$

d'où le niveau du test (SCCR) dans le cas où σ est inconnu :

$$\lim_{n\to\infty} -\frac{1}{n} \log P_{H_o}\{\exists\, t \in [0,1],\ \frac{s_{w^2,n}(t)}{s_{w^2,n}(1)} > t+g_1(t)\} = \frac{1}{2} \inf_{t\in[0,1]} \{t \log[1+\frac{g_1(t)}{t}]+(1-t)\log[(1-\frac{g_1(t)}{1-t})\vee 0]\}$$

Le calcul du niveau du test (SCR) dans le cas où σ est inconnu relève du même principe mais sa résolution est intéressante en soi parce que pathologique.

On veut évaluer :

$$\lim_{n\to\infty} -\frac{1}{n} \log P_{H_o}\{\exists\, t \in [0,1],\ \frac{s_{w,n}(t)}{\sqrt{s_{w^2,n}(1)}} > g(t)\}$$

où g est une fonction strictement positive, de classe C_1. Il est intéressant de remarquer que la simplification qui consiste à considérer le problème dans un espace de dimension 1 (en écrivant l'évènement sous la forme : $s_{w,n}(.) \in G_g$, où G_g désigne

$\{ f \in \mathscr{C}_{[0,1]}, f(0) = 0 ; \exists\, t \in [0,1] \text{ tel que } \dfrac{f(t)}{(\int_0^1 f'(t)^2 dt)^{1/2}} > g(t) \}$, est illusoire ;

en effet $\text{int}(G) = \emptyset$ et $\text{ad}(G)$ est un cône de sommet 0. En revanche, le problème se traite aisément dans un espace de dimension 2 en considérant la marche aléatoire $s_{(w,w^2),n}(.)$ et le borélien

$$G_g^2 = \{ f = (f_1, f_2) \in \mathscr{C}_{[0,1]} \times \mathscr{C}_{[0,1]},\ f(0) = 0 \text{ et } \exists\, t \in [0,1]\ \frac{f_1(t)}{\sqrt{f_2(1)}} > g(t) \} .$$

Dans ce cas, on a $W(f) = \int_0^1 h[f_1'(t), f_2'(t)]dt$ avec $h(x,y) = \dfrac{y-1-\log(y-x^2)}{2}$, ce qui donne :

$$\lim_{n\to\infty} -\frac{1}{n} \log P\{ s_{(w,w^2),n}(.) \in G_g^2 \} = \sup_{t\in[0,1]} \frac{1}{2} \log [(1 - \frac{g^2(t)}{t})] \vee 0] .$$

IV.3.c) Application à un non-franchissement de barrière.

Pour une suite de fonctions g_n absolument continues qui converge vers une fonction g absolument continue sur $[0,1]$ vérifiant :

$$\begin{cases} \exists\, \tau \in]0,1] \text{ tel que } g(\tau) < 0 \\ \exists\, K > \text{ess inf } \xi \text{ tel que } \dfrac{g_n(t)}{t} \geq K,\ \forall\, t \in]0,1], \end{cases}$$

on définit $G_{g_n} = \{ f \in \mathscr{C}_{[0,1]} \text{ telles que } \forall\, t \in [0,1],\ f(t) \leq g_n(t) \}$. Alors, $\lim_{n\to\infty} -\frac{1}{n} \log P\{ s_{\xi,n}(.) \in G_{g_n} \} = W(g^*)$, la fonction g^* constituant l'enveloppe convexe inférieure de l'ensemble $|(0,0) \cup (1, \inf_{t\in[0,1]} g(t)) \cup \{(t,y) \text{ tels que } y \geq g(t)\}|$.

La démonstration correspondant au cas particulier $g_n \equiv g$ figure dans [14]. De la convergence uniforme de g_n vers g, on déduit la convergence uniforme de g_n^* vers g^*, l'absolue continuité de g_n^* et g^* ; les $g_n'^*$ étant croissants à valeurs dans $[K,0]$, l'application successive du théorème de Helly et du théorème de convergence dominée fournit la convergence de $\int_0^1 h(g_n^{*\prime}(t))dt$ vers $\int_0^1 h(g^{*\prime}(t))dt$ et le résultat.

(De façon évidente, le résultat ci-dessus reste encore valable si $g(0) = 0$ mais il faut alors que $g_n(0)$ soit strictement positif pour tout n, ou bien que $g_n(0)$ soit nul mais alors on imposera que : $\lim_{n\to\infty} g_n'(0) = g'(0) \geq K$).

Les erreurs de 2^e espèce du test (SCR) sont respectivement évaluées par les expressions suivantes (σ connu) :

$$\begin{cases} \lim_{n\to\infty} -\frac{1}{n} \log P_{(\beta^*,\beta^{**},t_o)}\{ \forall\, t,\ \frac{s_{w,n}(t)}{\sigma} \le at + \frac{a}{2} \} \\ = t_o\Delta^2[t_o\varphi(\log\varphi + \frac{\varphi}{2} + 1 + \log t_o) + e^{-(\frac{\varphi}{2}+1)}] \quad \text{si} \quad 1 \le \varphi \le (t_o) \\ = t_o\Delta^2[t_o\varphi(\log\varphi + \frac{\varphi}{2} + 1 + \log t_o) + e^{-(\frac{\varphi}{2}+1)}] + \frac{t_o}{2}[\varphi^2 - \varphi\log\varphi - 1] \\ \qquad\qquad \text{si } \varphi \le \inf\{1, 2(\log\frac{1}{t_o} - 1)\} \\ = t_o^2\Delta^2[\frac{9}{8}\varphi^2 + \frac{3}{2}\varphi\log t_o + (\frac{\log t_o}{2})^2] \quad \text{si} \quad 2(\log\frac{1}{t_o} - 1) \le \varphi \le \frac{2}{3}\log\frac{1}{t_o} \\ = 0 \quad \text{sinon,} \end{cases}$$

où nous avons noté $\Delta = \frac{|\beta^* - \beta^{**}|}{\sigma}$, $\varphi = \frac{a}{t_o\Delta}$ et $\varphi(t_o)$ désigne la solution de l'équation :

$$\log t_o\,\varphi(t_o) + 1 + \frac{\varphi(t_o)}{2} = 0.$$

Pour évaluer l'erreur de 2^e espèce du test (SCCR) dans le cas où σ est connu, nous sommes amenés à traiter un problème de non franchissement de barrière pour des variables aléatoires non stationnaires. Il convient alors d'utiliser des résultats similaires démontrés par VENTZEL [18]. On obtient :

$$\lim_{n\to\infty} -\frac{1}{n} \log P_{H_1}(\forall\, t\ ,\ \frac{s_{w^2,n}(t)}{\sigma^2} \le t + c) = \inf_{f, f(t)\le t+c} \int_0^1 h_t[f'(t)]dt$$

où h_t désigne la transformée de Cramer de la variable $\frac{w^2}{2}[nt]$ sous l'hypothèse (β^*,β^{**},t_o). Un calcul de variation permet d'obtenir les expressions suivantes :

$$\begin{cases} \lim_{n\to\infty} -\frac{1}{n} \log P_{(\beta^*,\beta^{**},t_o)}\{ \forall\, t,\ s_{w^2,n}(t) \le t+c \} \\ = \frac{1}{2}[\log\Lambda - \frac{\Lambda-1}{\Lambda}(1 - \ell^2\frac{\Lambda-1}{\Lambda})] \text{ si } t_o < \frac{1}{2} \text{ et } \frac{4t_o(1-2t_o)\Delta^2}{(1+\sqrt{1+4t_o\Delta^2})^2} < c < \ell^2 \\ \qquad \text{avec } \Lambda = \frac{1+\sqrt{1+4(c+1)\ell^2}}{2(1+c)} \quad \text{ou si } t_o \ge \frac{1}{2} \\ = \frac{1}{2}\log\frac{1+\sqrt{1+4t_o^2\Delta^2}}{2} + 1 - \frac{t_o^2\Delta^2}{2} + t_o\Delta\,\frac{\Lambda^2-1}{\sqrt{\Lambda(\Lambda-1)}} - \sqrt{1+4t_o^2\Delta^2} + \frac{t_o\Delta^2}{2}(\frac{\Lambda-1}{\Lambda})^2 \\ \qquad + \frac{t_o\Delta}{2}\log\frac{2t_o\Delta+\sqrt{1+4t_o^2\Delta^2}}{2\sqrt{\Lambda(\Lambda-1)+(2\Lambda-1)}} \quad \text{si } t_o < \frac{1}{2} \text{ et } 0 < c < \frac{4t_o(1-2t_o)\Delta^2}{(1+\sqrt{1+4t_o^2\Delta^2})^2} \end{cases}$$

avec Λ solution de : $(\Lambda^2\frac{c}{t_o} - \Delta^2)^2 + 4\Delta\,\Lambda(\Lambda-1) = 0$.

En ce qui concerne le calcul des erreurs de 2^e espèce des tests (SCR) et (SCCR) dans le cas σ inconnu, les techniques employées sont identiques à la précédente mais les expressions obtenues sont trop compliquées pour avoir un intérêt pratique.

BIBLIOGRAPHIE

[1] *AZENCOTT R. "Grandes déviations". Cours de l'Ecole d'Eté de Probabilités de Saint-Flour (1978), Lectures Notes in Mathematics n° 774.*

[2] *BAHADUR R.R. "An optimal property of the likelihood ratio statistic" V^th Berkeley Symposium (1966), I, p. 12-26.*

[3] *BARTFAI P. "Connections between the convex analysis and the theory of large deviations" Preprint of the Mathematical Institute of Hungarian Academy of Sciences. (1977), n° 2.*

[4] *BOROVKOV A.A. "Boundary-value problems for random walks and large deviations in function spaces" T.P.A. (1967), n° 12, p. 575-595.*

[5] *BROWN L.D. "Non-local asymptotic optimality of appropriate likehood ratio tests" A.M.S. (1971) n° 42, p. 1206-1240.*

[6] *BROWN R.L. DURBIN J. EVANS J.M. "Techniques for testing the constancy of regression relationships over time". J.R.S.S. (1975) B, p. 149-195.*

[7] *CHERNOFF H. "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations" A.M.S. (1952) n° 23, p. 493-507.*

[8] *DACUNHA-CASTELLE D. "Formule de Chernoff pour une suite de variables réelles". Séminaire d'Orsay (1977-1978). Astérisque n° 68, p.19-24.*

[9] *FEDER P.I. "On asymptotic distribution theory in segmented regression problems Identified case". A.S. (1975) n° 3, p. 49-83.*

[10] *FEDER P.I. "The log-likehood ration in segmented regression". A.S. (1975) n° 3, p. 84-97.*

[11] *HAWKINS D.M. "Testing a sequence of observations for a shift in location" J.A.S.A. (1977) n° 72, p. 180-186.*

[12] *HINKLEY D.V. "Inference about the change-point in a sequence of random variables". Biometrika (1970) n° 57, p. 1-17.*

[13] *KRANOSEL'SKII RUTICKII. "Convex functions and Orlicz spaces". Noordhoff Ltd.*

[14] *MOGULSKII. "Large deviations for trajectories of multidimensional random walks" T.P.A. (1976) n° 21, p. 300-315.*

[15] *QUANDT R.E. "Tests of the hypothesis that a linear regression system obeys two separate regimes". J.A.S.A. (1960) n° 55, p. 324-330.*

[16] *SEN A. SRIVASTAVA M.S. "On tests for detecting change in mean". A.S. (1975) n° 3, p. 98-108.*

[17] *TUSNADY G. "On asymptotically optimal tests". A.S. (1977) n° 5, p. 385-393.*

[18] *VENTZEL A.D. "Rough limit theorems on large deviations for Markov processes". T.P.A. (1976) n° 21, I, p. 227-242 et II, p. 499-512.*

Université Paris-Sud
ERA C.N.R.S. 532 "Statistique Appliquée"
Bât. de Mathématiques 425

91405 ORSAY Cedex (France)

PRINCIPE D'INVARIANCE SUR LE PROCESSUS DE VRAISEMBLANCE

Le but de cet article est de démontrer un principe d'invariance sur le processus de vraisemblance. Le résultat le plus précis existant sur le comportement asymptotique classique du processus de vraisemblance ([5], [7]) démontre la convergence étroite de celui-ci vers un processus gaussien dégénéré dans l'espace $\mathcal{E}_o(\mathbb{R}^p)$ lorsque le paramètre est p-dimensionnel sous des conditions de régularité relativement faibles. Un des intérêts de la généralisation de l'invariance réside dans la nécessité d'élargir l'espace $\mathcal{E}_o$ qui n'est plus adapté à cause des problème de bords. Ce type de résultats est indispensable pour certaines applications statistiques : celle traitée dans [12] concerne les problèmes de détection de rupture dans un modèle.

The aim of this paper is to prove an invariance principle on likelihood process. The very accurate result ([5], [7]) on classical asymptotic behaviour of the likelihood process claims the weak convergence to a degenerated gaussian process in the space $\mathcal{E}_o(\mathbb{R}^p)$ when the parameter is p-dimensional under weak regularity conditions. One of the interest of the generalization to invariance principale is that the space $\mathcal{E}_o$ is no nore suitable and must be extended because of edges-problems. This kind of results is essential for some applications in statistics as those investigated in [12] about change point detection.

INTRODUCTION

Dans les problèmes de détection de ruptures de modèle en statistique (problèmes de tests ou d'estimation des ruptures et de leurs amplitudes dans un modèle que l'on sait à plusieurs phases), la recherche de procédures asymptotiquement optimales et le comportement des techniques associées à la vraisemblance reposent sur un théorème d'invariance sur la suite des processus de vraisemblance : celui-ci met en évidence la nécessité d'une renormalisation des paramètres dans le test du rapport de vraisemblance, et fournit des lois limites conjointes des estimateurs des amplitudes et des instants de rupture qui ne sont pas classiques (voir [11], [12]).

Le résultat d'invariance sous les conditions le s plus larges possibles dépasse le domaine d'application des problèmes de rupture et présente un intérêt en soi : la convergence classique du processus de vraisemblance dans $\mathcal{C}_o(\mathbb{R}^p)$ lorsque le paramètre est p-dimensionnel, est obtenue par IBRAGIMOV-KHASMINSKII ([5],[6]) sous des hypothèses relativement faibles sur le modèle ; la généralisation par invariance ici est non classique en ce sens que sous des conditions équivalentes, la convergence n'est plus vraie dans le même espace $\mathcal{C}_o(\mathbb{R}^p \times [0,1])$; il est alors nécessaire d'agrandir l'espace et d'introduire $\mathcal{C}_\varphi$ pour évaluer la vitesse de convergence au voisinage du point 0 de l'intervalle $[0,1]$.

Le premier paragraphe de cet article est consacré à la convergence dans $\mathcal{C}(\mathbb{R}^p \times [0,1])$ muni de la topologie de la convergence uniforme sur les compacts (théorème d'invariance "faible"). Le second paragraphe étudie la convergence dans l'espace $\mathcal{C}_\varphi(\mathbb{R}^p \times [0,1])$ que nous allons définir dans un instant (théorème d'invariance "fort") ; à la fin de la deuxième partie on examine des élargissement possibles des hypothèses posées par IBRAGIMOV-KHASMINSKII [6].

Précisons maintenant le modèle et les notations. Soient $X_1,\ldots,X_n$ n variables aléatoires indépendantes à valeurs dans $\mathcal{X}$, espace polonais de lois $(P_\theta)_{\theta\in\Theta}$ absolument continues par rapport à une mesure σ-finie μ, de densité $g(\cdot,\theta)$, l'ensemble Θ étant identifié à un ouvert de l'espace euclidien $\mathbb{R}^p$ muni de sa norme $\|\cdot\|$. On considère le processus de vraisemblance $Z_n(\cdot,\cdot)$ à valeurs dans $\mathcal{C}(\mathbb{R}^p \times [0,1])$: espace des fonctions continues sur $\mathbb{R}^p \times [0,1]$, défini comme suit :

pour tout θ appartenant à $\mathbb{R}^p$, $Z_n(\theta,\cdot)$ est la ligne polygonale joignant les points $(0;1)$ et $\left(\frac{k}{n}, \left[\prod_{i=1}^{k} \frac{g(X_i,\theta_o+\frac{\theta}{\sqrt{n}})}{g(X_i,\theta_o)}\right]^{\frac{1}{q}}\right)$ $k = 1,2,\ldots,n$.

Le point θ_o est la "vraie valeur" du paramètre et les autres points correspondent aux autres valeurs possibles du paramètre, renormalisées autour de θ_o pour que la suite d'expériences indicée par n converge lorsque le modèle est régulier. Le nombre positif q n'a d'intérêt que technique ; sa valeur (déterminée par les hypothèses sur les lois P_θ) ne sera fixée que plus tard. Nous montrons la convergence de la suite de processus $Z_n(\cdot,\cdot)$ dans un sous-espace $\mathcal{C}_\varphi$ de $\mathcal{C}(\mathbb{R}^p \times [0,1])$ défini comme suit. Soit φ une fonction continue sur $[0,1]$, nulle en 0 telle que $\frac{\varphi(t)}{\sqrt{t}}$ soit croissante sur $[0,1]$ et valant 1 en 1. Nous appelerons $\mathcal{C}_\varphi$ l'ensemble des fonctions z continues sur $\mathbb{R}^p \times [0,1]$, valant 1 sur $\mathbb{R}^p \times \{0\}$ et telles que $\lim_{M\to\infty} \sup_t \sup_{\varphi(t)\|\theta\|\geq M} |z(\theta,t)| = 0$. Cet ensemble

$\mathcal{C}$ est muni de la topologie définie comme suit :

on dira que z_n converge vers z si et seulement si :

i) pour tout ε positif, pour tout compact K de $\mathbb{R}^p$, il existe $n_o(\varepsilon,K)$ tel que pour tout n supérieur à n_o, $\sup_{(\theta,t)\in K\times[0,1]} |z_n(\theta,t)-z(\theta,t)| \leq \varepsilon$.

ii) pour tout ε positif, il existe $M_o(\varepsilon)$ tel que pour tout M supérieur à M_o et pour tout n , $\sup_{\substack{(\theta,t)\\ \varphi(t)||\theta||\geq M}} |z_n(\theta,t)| \leq \varepsilon$

Nous ne détaillerons pas plus cette topologie qui est d'ailleurs métrisable par la distance :

$$d(z_1,z_2) = \sum_{m=o}^{\infty} \frac{1}{2^m} \left[\sup_{\substack{(\theta,t)\\ m\leq||\theta||<m+1}} |z_1(\theta,t)-z_2(\theta,t)| \wedge 1 \right] + \lim_{M\to\infty} \sup_{\substack{(\theta,t)\\ \varphi(t)||\theta||>M}} |z_1(\theta,t)-z_2(\theta,t)|$$

Nous nous bornerons à en décrire les ensembles relativement compacts ; d'après le théorème d'Ascoli, un ensemble $\mathcal{Z}$ inclus dans $\mathcal{C}_\varphi$ est relativement compact si et seulement si :

i) pour tout ε positif, pour tout compact K , il existe $\delta(\varepsilon,K)$ tel que :

$$\sup_{z\in\mathcal{Z}} \sup_{\substack{(\theta_1,t_1)\in K\times[0,1]\\ (\theta_2,t_2)\in K\times[0,1]\\ ||(\theta_1,t_1)-(\theta_2,t_2)||\leq\delta}} |z(\theta_1,t_1) - z(\theta_2,t_2)| \leq \varepsilon$$

ii) pour tout ε positif, il existe $M(\varepsilon)$ tel que :

$$\sup_{z\in\mathcal{Z}} \sup_{\substack{(\theta,t)\\ \varphi(t)||\theta||\geq M}} |z(\theta,t)| \leq \varepsilon \ .$$

I - Convergence de la suite de processus Z_n dans $\mathcal{C}(\mathbb{R}^p \times [0,1])$

Dans ce paragraphe, nous formulerons, pour simplifier, les hypothèses sur la suite de processus Z_n ; au paragraphe II, diverses conditions suffisantes sur le modèle seront explicitées assurant que ces hypothèses sont vérifiées.

A1) la matrice d'information de Fisher existe au point θ_o et est définie positive :

$$I(\theta_o) = \int_{\mathcal{X}} \frac{\partial}{\partial\theta} g^{\frac{1}{2}}(x,\theta_o) \, {}^t\frac{\partial}{\partial\theta} g^{\frac{1}{2}}(x,\theta_o) d\mu(x)$$

où $\frac{\partial}{\partial\theta} g^{\frac{1}{2}}(x,\theta_o)$ désigne le gradient (en θ) de $g^{\frac{1}{2}}(x,\cdot)$ au point θ_o , et les marginales de dimension finie de la suite de processus $Z_n(\cdot;1)$ de $\mathcal{C}(\mathbb{R}^p)$ convergent vers celles du processus $Z(\cdot,1)$ défini par :

$$Z(\theta,1) = \exp \frac{1}{q} \{ {}^t\theta \, I(\theta_o)^{\frac{1}{2}} \, \xi - \frac{1}{2} \, {}^t\theta \, I(\theta_o)\theta \}$$

où ξ est un vecteur aléatoire gaussien de dimension p, centré, de covariance identité.

B1) la suite de processus $Z_n(\cdot\,;1)$ de $\mathcal{C}(\mathbb{R}^p)$ vérifie une condition de "mesure majorante sur les pavés" (que nous expliciterons plus loin après quelques définitions et commentaires).

Théorème 1 : Sous les conditions A1 et B1, la suite de processus $Z_n(\cdot,\cdot)$ converge étroitement sous P_{θ_o} , dans $\mathcal{C}(\mathbb{R}^p \times [0,1])$ muni de la topologie de la convergence uniforme sur les compacts vers le processus Z :

$$Z(\theta,t) = \exp \frac{1}{q} \{ {}^t\theta\, I(\theta_o)^{\frac{1}{2}}\, W(t) - \frac{t}{2} \cdot {}^t\theta\, I(\theta_o)\theta \}$$

où $W(\cdot)$ est un mouvement Brownien réel à p dimensions sur $[0,1]$.

Remarque : Dans le cas où Θ n'est pas l'espace $\mathbb{R}^p$ tout entier, le processus Z_n n'est défini à priori que sur l'ensemble $\mathcal{D}_n = \{\theta,\ \theta_o + \frac{\theta}{\sqrt{n}} \in \Theta\}$ mais il est facile de poser :

$$Z_n(\theta\,;0) \equiv 1\ , \quad \forall\theta \in \mathbb{R}^p$$

$Z_n(\theta, \frac{k}{n}) = 0$ pour θ appartenant au complémentaire d'un voisinage de $\mathcal{D}_n$, $k = 1,\ldots,n$ et de le prolonger en un processus continu sur $\mathbb{R}^p$ tout entier sans changer les propriétés de convergence ; désormais, nous supposerons $\Theta = \mathbb{R}^p$.

Démonstration du théorème 1

La convergence des marginales de dimension finie de la suite $Z_n(\cdot,\cdot)$ est évidente à cause de l'indépendance des variables X_i et de l'hypothèse A1. La tension résultera des deux lemmes suivants :

Lemme 1.1. : Pour tout compact K de $\mathbb{R}^p$, et pour tout ε positif, il existe $\delta(\varepsilon,K)$ positif tel que pour tout n ,

$$P\{ \sup_{\substack{t_1,t_2\in[0,1] \\ |t_1-t_2|\leq\delta}} \sup_{\theta\in K} |Z_n(\theta,t_1) - Z_n(\theta,t_2)| \geqslant \varepsilon\} \leqslant \varepsilon \qquad (1.1)$$

Lemme 1.2. : Pour tout compact K de $\mathbb{R}^p$ et pour tout ε positif, il existe $\delta(\varepsilon,K)$ positif tel que pour tout n

$$P\{ \sup_{t\in[0,1]} \sup_{\substack{\theta_1,\theta_2\in K \\ \|\theta_1-\theta_2\|\leq\delta}} |Z_n(\theta_1,t)-Z_n(\theta_2,t)| \geqslant \varepsilon\} \leqslant \varepsilon \qquad (1.2)$$

Pour démontrer ces deux lemmes, nous utiliserons le critère de CSENSOV [4] qui généralise en dimension p le critère de tension de KOLMOGOROV. Une généralisation possible du lemme de KOLMOGOROV (cf. théorème 12.2. de BILLINGSLEY [2]) consiste à assimiler le processus Y à une fonction de répartition sur $\mathbb{R}$: $Y(\theta_2) - Y(\theta_1)$ et à considérer dans $\mathbb{R}^p$: $Y[\theta_1,\theta_2]$ comme la mesure du processus Y sur le pavé $[\theta_1,\theta_2]$ définie au moyen des notations suivantes :

sur $\mathbb{R}^p$, nous définissons la relation d'ordre (partiel) entre θ_1 et θ_2 :

$\theta_1 \leqslant \theta_2$ si et seulement si pour toutes les coordonnées $i = 1,\ldots,p$, on a : $\theta_1^{(i)} \leqslant \theta_2^{(i)}$; dans ce cas, le pavé $[\theta_1, \theta_2]$ est défini sans ambiguité. Si Y est un processus défini sur $\mathbb{R}^p$, nous allons définir le "processus de pavé" $Y[\theta_1,\theta_2]$ pour $\theta_1 \leqslant \theta_2$, par récurrence sur la dimension p. Le cas $p = 1$ étant déjà vu au départ, considérons $p > 1$ et introduisons tous les sous-ensembles u non vides de $\{1,2,\ldots,p\}$ de cardinal $\ell(u) \leqslant p$; notons $\bar{u}$ le complémentaire de u et π_u la projection de $\mathbb{R}^p$ dans $\mathbb{R}^{\ell(u)}$ qui ne conserve que les coordonnées à indices dans u ; on notera souvent θ^u pour $\pi_u(\theta)$. Pour θ appartenant à $\mathbb{R}^p$, le processus $Y(\theta^{\bar{u}},\cdot)$ est défini sur $\mathbb{R}^{\ell(u)}$ par $Y(\theta^{\bar{u}},\lambda) = Y((\theta^{\bar{u}},\lambda))$ où le vecteur $(\theta^{\bar{u}},\lambda)$ de $\mathbb{R}^p$ a ses coordonnées réordonnées de sorte que celles de λ correspondent aux indices dans u. Supposons que le processus de pavé sur $\mathbb{R}^{\ell(u)}$ ait été défini par les étapes inférieures de la récurrence ($\ell(u) < p$) : $Y(\theta^{\bar{u}}, [\lambda_1,\lambda_2])$, on peut alors définir le processus de pavé sur $\mathbb{R}^p$ par :

$$Y[\theta_1,\theta_2] = Y(\theta_2) - Y(\theta_1) + \sum_{\substack{u \\ 1\leqslant\ell(u)<p}} (-1)^{p-\ell(u)}\, Y(\theta_1^{\bar{u}}, \Pi_u[\theta_1,\theta_2]) \qquad (1.3)$$

Lemme 1.3. (CSENSOV) : Soit Y un processus de $\mathcal{E}([0,1]^p)$ tel qu'il existe des constantes $B > 0$; $\gamma > 0$ et $\alpha > 1$ et une mesure ν finie sur $[0,1]^p$ vérifiant $\forall\theta_1, \forall\theta_2 \in [0,1]^p$, $\theta_1 \leqslant \theta_2$, $\forall A > 0$:

$$P\{|Y[\theta_1,\theta_2]| > A\} \leqslant \frac{B}{A^\gamma} (\nu[\theta_1,\theta_2])^\alpha \qquad (1.4)$$

Alors il existe une constante C_p positive (ne dépendant que de la dimension p) telle que : $\forall\underline{\theta}$, $\forall\overline{\theta} \in [0,1]^p$, $\underline{\theta} \leqslant \overline{\theta}$, $\forall A > 0$

$$P\{\sup_{\substack{\theta_1,\theta_2 \\ [\theta_1,\theta_2]\subseteq[\underline{\theta},\overline{\theta}]}} |Y[\theta_1,\theta_2]| > A\} < \frac{C_p.B}{A^\gamma} (\nu[\underline{\theta},\overline{\theta}])^\alpha$$

Nous donnons une démonstration de ce lemme par récurrence sur la dimension p. Elle précise celle de BICKEL-WICHURA [1] par l'usage indispensable en dimension $p > 2$ d'un théorème de section, omis dans [1] et généralise une démonstration de YOR [10] (en dimension 2). De plus, au fur et à mesure de cette démonstration, nous introduisons les notations reprises dans la suite.

Lemme 1.4. : Soit $\tilde{K}$ un ensemble compact et $\tilde{Y}$ un processus de $\mathcal{E}(\tilde{K})$ à valeurs dans $\mathbb{R}^r$. Alors il existe une variable aléatoire $\hat{\lambda}_{\tilde{K}}$ à valeurs dans $\tilde{K}$ telle que :

$$\sup_{\lambda\in\tilde{K}} ||\tilde{Y}(\lambda)||_r = ||Y(\hat{\lambda}_{\tilde{K}})||_r \quad .$$

La démonstration est classique à partir du théorème de section mesurable (PARTHASARATY [8]).

Démonstration du lemme 1.3.

La récurrence est amorcée sur la dimension p est amorcée avec $p = 1$: c'est exactement le théorème 12.2 de BILLINGSLEY [2] .

Supposons donc p strictement supérieur à 1 et que la conclusion du lemme est vraie en dimension $p-1$. Nous prenons $u = \{1,2,\ldots,p-1\}$ et donc $\bar{u} = \{p\}$. Le processus

$\tilde{Y}(\lambda_1,\lambda_2) = \sup_{\underline{\theta}^{(p)} \leq \theta_1^{(p)} \leq \theta_2^{(p)} \leq \bar{\theta}^{(p)}} Y\left[(\lambda_1,\theta_1^{(p)}),(\lambda_2,\theta_2^{(p)})\right]$ étant continu sur le sous-ensemble compact de $\mathbb{R}^{p-1} \times \mathbb{R}^{p-1}$: $\underline{\theta}^u \leq \lambda_1 \leq \lambda_2 \leq \bar{\theta}^u$, le lemme 1.4. assure l'existence de deux variables aléatoires $\hat{\lambda}_1$ et $\hat{\lambda}_2$ safisfaisant $\underline{\theta}^u \leq \hat{\lambda}_1 \leq \hat{\lambda}_2 \leq \bar{\theta}^u$ et telles que :

$$\sup_{\substack{\theta_1,\theta_2 \\ \theta_1 \leq \theta_2,\ [\theta_1,\theta_2] \subseteq [\underline{\theta},\bar{\theta}]}} |Y[\theta_1,\theta_2]| = \sup_{\underline{\theta}^{(p)} \leq \theta_1^{(p)} \leq \theta_2^{(p)} \leq \bar{\theta}^{(p)}} |Y\left[(\hat{\lambda}_1,\theta_1^{(p)}),(\hat{\lambda}_2,\theta_2^{(p)})\right]|$$

En remarquant que :

$$Y\left[(\hat{\lambda}_1,\theta_1^{(p)}),(\hat{\lambda}_2,\theta_2^{(p)})\right] = Y(\theta_2^{(p)} ; [\hat{\lambda}_1,\hat{\lambda}_2]) - Y(\theta_1^{(p)} ; [\hat{\lambda}_1,\hat{\lambda}_2])$$

il suffira pour prouver le lemme 1.3. en dimension p, de montrer en utilisant le théorème 12.2. de BILLINGSLEY [2] que le processus à un indice $(\theta^{(p)})$: $Y(\theta^{(p)}, [\hat{\lambda}_1,\hat{\lambda}_2])$ vérifie :

$$P\{|Y(\theta_2^{(p)}, [\hat{\lambda}_1,\hat{\lambda}_2]) - Y(\theta_1^{(p)}, [\hat{\lambda}_1,\hat{\lambda}_2])| > A\} \leq \frac{B_1}{A^\gamma} \left[F(\theta_2^{(p)}) - F(\theta_1^{(p)})\right]^\alpha$$

avec $\alpha > 1$ et F croissante sur $[\underline{\theta}^{(p)}, \bar{\theta}^{(p)}]$ pour en déduire que :

$$P\{\sup_{\underline{\theta}^{(p)} \leq \theta_1^{(p)} \leq \theta_2^{(p)} \leq \bar{\theta}^{(p)}} |Y\left[(\hat{\lambda}_1,\theta_1^{(p)}),(\hat{\lambda}_2,\theta_2^{(p)})\right] > A\} \leq \frac{C_1 \cdot B_1}{A^\gamma} \left[F(\bar{\theta}^{(p)}) - F(\underline{\theta}^{(p)})\right]^\alpha$$

Or il est facile de voir que :

$$P\{|Y(\theta_2^{(p)}, [\hat{\lambda}_1, \hat{\lambda}_2]) - Y(\theta_1^{(p)}, [\hat{\lambda}_1, \hat{\lambda}_2]| > A\}$$

$$\leq P\{ \sup_{\substack{\lambda_1, \lambda_2 \\ [\lambda_1, \lambda_2] \subseteq [\underline{\theta}^u, \overline{\theta}^u]}} |Y[(\lambda_1, \theta_1^{(p)}), (\lambda_2, \theta_2^{(p)})]| > A\} \leq \frac{C_{p-1} \cdot B}{A^{\gamma}} (\nu[(\underline{\theta}^u, \theta_1^{(p)})(\overline{\theta}^u, \theta_2^{(p)})])^{\alpha}$$

en appliquant la propriété de récurrence en dimension p-1 ; ce qui achève la preuve.

On est maintenant en mesure d'en déduire un critère de relative étroite compacité dans $\mathcal{C}(\mathbb{R}^p)$ muni de la topologie de la convergence uniforme sur les compacts.

Critère de CSENSOV :

Une famille $(Y_n)_{n \in \mathcal{N}}$ de processus de $\mathcal{C}(\mathbb{R}^p)$ est tendue dès qu'il existe un point θ_o de $\mathbb{R}^p$ pour lequel

i) la famille $Y_n(\theta_o)$ est tendue

ii) pour tout $u \subseteq \{1,2,\ldots,p\}$, il existe des constantes $\gamma_u > 0$, $\alpha_u > 1$, une mesure ν_u de Radon sur $\mathbb{R}^{\ell(u)}$ à marginales diffuses et pour tout compact K , il existe des constantes $B_u(K)$ satisfaisant : $\forall n \in \mathcal{N}$, $\forall \theta_1, \forall \theta_2 \in K$, $\theta_1 \leq \theta_2$, $\forall A > 0$,

$$P\{|Y_n(\theta_o^{\overline{u}}, \Pi_u[\theta_1, \theta_2])| > A\} \leq \frac{B_u(K)}{A^{\gamma_u}} \{\nu_u(\Pi_u[\theta_1, \theta_2])\}^{\alpha_u}$$

Démonstration du critère

Quitte à agrandir le compact K, on peut supposer qu'il contient θ_o qui sera pris comme origine de $\mathbb{R}^p$: $\theta_o = 0$.

Le processus Y_n se décompose en la somme de 2^p processus :

$$Y_n(\theta) = \sum_{u \subseteq \{1,2,\ldots,p\}} Y_n(0^{\bar{u}}, \pi_u[0,\theta]) + Y_n(0)$$

et il suffit de montrer que chacun est tendu dans $\mathcal{C}(\mathbb{R}^p)$; en fait cela revient à montrer que $Y_n(0^{\bar{u}}, \pi_u[0,\theta])$ forme une famille tendue dans $\mathcal{C}(\mathbb{R}^{\ell(u)})$.

Tous les indices jouant un rôle symétrique, nous allons étudier un seul processus de face $Y_n(0^{\bar{u}}, \pi_u[0,\theta])$ avec $u = \{1,2,\ldots,\ell\}$, $1 \leq \ell \leq p$. Nous constatons que la démonstration est la même, quelque soit la dimension ℓ de la face passant par 0 et pour éviter d'écrire $0^{\bar{u}}$, nous le montrons pour $\ell = p$ sachant que pour un u quelconque, la mesure ν_u joue le même rôle que ν dans ce qui suit.

La tension du processus $Y_n[0,\theta]$ sur le compact K résultera du fait que pour chaque coordonnée $i = 1,2,\ldots,p$, on a :

$$\forall \varepsilon > 0, \ \exists \delta(\varepsilon,K),$$

$$P\{ \sup_{\substack{\theta_1^{(i)},\theta_2^{(i)} \in \pi_i(K) \\ |\theta_2^{(i)}-\theta_1^{(i)}|<\delta}} \sup_{\theta \in K} |Y_n[0,(\theta_2^{(i)},\theta^{(\bar{i})})] - Y_n[0,(\theta_1^{(i)},\theta^{(\bar{i})})]| \geq \varepsilon\} \leq \varepsilon$$

La démonstration en est classique et se fait en découpant le domaine réel $\pi_i(K)$ en intervalles de longueur δ et d'extrêmités $\tilde{\theta}_j^{(i)}$ indicés par $j \in J_i(K)$; on est ainsi ramené à majorer :

$$\sum_{j \in J_i(K)} P\{ \sup_{\tilde{\theta}_j^{(i)} \leq \theta^{(i)} \leq \tilde{\theta}_{j+1}^{(i)}} \sup_{\theta \in K} |Y_n[\tilde{\theta}_j^{(i)},\theta^{(\bar{i})}),(\theta^{(i)},\theta^{(\bar{i})})]| \geq \frac{\varepsilon}{3}\}$$

$$\leq C_p . B(K) . (\frac{3}{\varepsilon})^{\gamma} \sum_{j \in J_i(K)} \{\nu([\tilde{\theta}_j^{(i)},\tilde{\theta}_{j+1}^{(i)}] \times \pi_{\bar{i}}(K))\}^{\alpha}$$

$$\leq C_p . B(K) . (\frac{3}{\varepsilon})^{\gamma} \nu(K) . \sup_{j \in J_i(K)} \{\nu([\tilde{\theta}_j^{(i)},\theta_{j+1}^{(i)}] \times \pi_{\bar{i}}(K))\}^{\alpha-1}$$

les mesures ν_u étant à marginales diffuses, le critère est démontré.

Il nous est maintenant facile de préciser l'hypothèse B1) du théorème 1 laissée en suspens :

B1) la suite de processus $Z_n(\cdot;1)$ satisfait l'hypothèse ii) du critère de CSENSOV énoncé précédemment.

Remarque 1 : L'écriture $Y|\theta_1,\theta_2| = \int\cdots\int_{[\theta_1,\theta_2]} \frac{\partial^{(p)}}{\partial\theta^{(1)}\ldots\partial\theta^{(p)}} Y(v^{(1)},\ldots,v^{(p)})$ $dv^{(1)},\ldots,dv^{(p)}$ implique qu'un processus vérifiant pour un $\gamma > 1$ et pour tout compact K

$$E \sup_{\theta\in K} \left|\frac{\partial^p}{\partial\theta^{(1)}\ldots\partial\theta^{(p)}} Y(\theta)\right|^\gamma < \infty$$

satisfait les conditions du critère, avec ν la mesure de Lebesgue. En particulier, le processus $Y(\cdot) = Z(\cdot;1)$ où Z est le processus limite défini au théorème 1, satisfait cette dernière condition.

Remarque 2 : On constate ainsi que les hypothèses A1) et B1) impliquent immédiatement la convergence étroite dans $\mathcal{C}(\mathbb{R}^p)$ de la suite de processus $Z_n(\cdot,1)$ vers le processus $Z(\cdot,1)$. Toutefois, il nous est impossible d'utiliser tel quel le critère ci-dessus pour démontrer le théorème 1 à cause de la condition $\alpha > 1$, qui si elle est tout à fait adaptée vis à vis de la variable θ, ne l'est plus du tout en ce qui concerne la variable t : celle-ci vérifiant une condition du même type avec $\alpha = 1$. Il convient alors d'utiliser la structure d'accroissements indépendants (par rapport à t) du logarithme du processus. A cet effet nous utilisons le lemme classique suivant :

Lemme 1.5. : Soient $(\xi_1(\cdot), \ldots, \xi_n(\cdot)$ des fonctions aléatoires indépendantes, définies et continues sur un compact K de $\mathbb{R}^p$, à valeurs dans $\mathbb{R}^r$. Posons $S_k(\theta) = \sum_{j=1}^{k} \xi_j(\theta)$, et $\tau_A = \inf\{j, \sup_{\theta\in K} |S_j(\theta)|_r > A\}$. Alors pour tout A positif :

$$P(\tau_A \leqslant n) \leqslant P\{\sup_{\theta\in K} |S_n(\theta)|_r > \frac{A}{2}\} \left[1 - \sup_{k=1,\ldots,n-1} P\{\sup_{\theta\in K} |S_n(\theta)-S_k(\theta)|_r > \frac{A}{2}\}\right]^{-1}$$

Démonstration du lemme 1.5.

Il est facile de vérifier que :

$$P(\tau_A \leqslant n) \leqslant P\{\sup_{\theta\in K} |S_n(\theta)| > \frac{A}{2}\} + \sum_{k=1}^{n-1} P(\tau_A = k \cap \sup_{\theta\in K} |S_n(\theta)| \leqslant \frac{A}{2})$$

$$\leqslant P\{\sup_{\theta\in K} |S_n(\theta)| > \frac{A}{2}\} + \sum_{k=1}^{n-1} P(\tau_A = k \cap \sup_{\theta\in K} |S_n(\theta)-S_k(\theta)| > \frac{A}{2})$$

$$\leqslant P\{\sup_{\theta\in K} |S_n(\theta)| > \frac{A}{2}\} + \sum_{k=1}^{n-1} P(\tau_A = k) \cdot \sup_{k=1,\ldots,n-1} P\{\sup_{\theta\in K} |S_n(\theta)-S_k(\theta)| > \frac{A}{2}\}$$

$$\leqslant P\{\sup_{\theta\in K} |S_n(\theta)| > \frac{A}{2}\} + P(\tau_A \leqslant n) \cdot \sup_{k=1,\ldots,n-1} P\{\sup_{\theta\in K} |S_n(\theta)-S_k(\theta)| > \frac{A}{2}\}$$

ce qui conduit au résultat annoncé.

Démonstration du lemme 1.1.

Pour ε positif arbitrairement petit et δ positif (δ sera fixé plus tard en fonction de ε et du compact K, que nous prenons pour simplifier égal à $[0,D]^p$), nous découpons l'intervalle $[0,1]$ en intervalles de longueur δ et d'extrémités $\tau_{n,j} = \frac{j[n\delta]}{n}$, nous pouvons écrire la majoration :

$$P\{ \sup_{\substack{t_1,t_2\in[0,1] \\ |t_2-t_1|\leq\delta}} \sup_{\theta\in K} |Z_n(\theta,t_2) - Z_n(\theta,t_1)| > \varepsilon\}$$

$$\leq P\{ \sup_{j=0,\ldots,\left[\frac{1}{\delta}\right]} \sup_{\tau_{n,j}\leq t\leq\tau_{n,j+1}} \sup_{\theta\in K} |Z_n(\theta,t) - Z_n(\theta,\tau_{n,j})| > \frac{\varepsilon}{3}\}$$

$$\leq P\{ \sup_{j=0,\ldots,\left[\frac{1}{\delta}\right]} \sup_{\theta\in K} Z_n(\theta,\tau_{n,j}) > A\}$$

$$+ \sum_{j=0}^{\left[\frac{1}{\delta}\right]} P\{ \sup_{\tau_{n,j}\leq t\leq\tau_{n,j+1}} \left|\frac{Z_n(\theta,t)}{Z_n(\theta,\tau_{n,j})} - 1\right| > \frac{\varepsilon}{3A}\}$$

où $A(\varepsilon,D)$ sera fixé ci-dessous $(A > 1)$

$$\leq P\{ \sup_{j=0,\ldots,\left[\frac{1}{\delta}\right]} \sup_{\theta\in K} \text{Log } Z_n(\theta, \frac{j\left[n\delta\right]}{n}) \geq \text{Log } A\}$$

$$+ \frac{1}{\delta} P\{ \sup_{k=1,\ldots,\left[n\delta\right]} \sup_{\theta\in K} |Z_n(\theta,\frac{k}{n}) - 1| \geq \frac{\varepsilon}{3A}\} = \pi_1 + \pi_2 .$$

Chacun des deux termes peut être majoré séparemment en utilisant le lemme 1.5. :

$$\pi_1 \leq P\{\sup_{\theta\in K} |\text{Log } Z_n(\theta;1)| \geq \frac{\text{Log } A}{2}\} \times$$

$$\left[1 - \sup_{j=1,\ldots,\left[\frac{1}{\delta}\right]} P\{\sup_{\theta\in K} |\text{Log } Z_n(\theta, \frac{j\left[n\delta\right]}{n})| > \frac{\text{Log } A}{2}\right]^{-1}$$

en utilisant la "stationnarité en t" du processus Z_n : les variables X_i sont indépendantes, identiquement distribuées.

La remarque 2 suivant la démonstration du critère de Csensov affirme que la suite de processus $Z_n(\cdot;1)$ converge vers $Z(\cdot,1)$ uniformément sur tout compact ; d'autre part, le processus limite étant strictement positif presque sûrement, on en déduit que la suite $\text{Log } Z_n(\cdot;1)$ converge vers $\text{Log } Z(\cdot,1)$ dans $\mathcal{C}(\mathbb{R}^p)$ et par suite :

$$P\{\sup_{\theta\in K} |\text{Log } Z_n(\theta,1)| \geq \frac{\text{Log } A}{2}\} \leq 2P\{\sup_{\theta\in K} |\text{Log } Z(\theta,1)| \geq \frac{\text{Log } A}{2}\} \leq \frac{\varepsilon}{4} \quad (1.8)$$

pour $A \geq A_o(K,\varepsilon)$ et pour $n \geq n_o(\varepsilon,K,A_o)$; la dernière inégalité est facilement réalisée en utilisant le moment d'ordre 2 de ξ .

De l'inégalité (1.8), on déduit aussi une minoration du dénominateur de π_1 par $1 - \frac{\varepsilon}{4}$ dès que $[n\delta] \geq n_o(\varepsilon,K,A_o)$; soit :

$$\pi_1 \leq \frac{\varepsilon}{4-\varepsilon} \leq \frac{\varepsilon}{2} \quad \text{pour} \quad n \geq \frac{n_o(\varepsilon,K,A_o)}{\delta} \quad .$$

La majoration de π_2 à l'aide du lemme 1.5. conduit de même à :

$$\pi_2 \leq \frac{1}{\delta} P\{\sup_{k=1,\ldots,[n\delta]} \sup_{\theta\in K} |\text{Log } Z_n(\theta,\tfrac{k}{n})| > \frac{\varepsilon}{6A_o}\}$$

$$\leq \frac{1}{\delta} P\{\sup_{\theta\in K} |\text{Log } Z_n(\theta, \frac{[n\delta]}{n})| \geq \frac{\varepsilon}{12A_o}\} \times$$

$$\left[1 - \sup_{k=1,\ldots,[n\delta]} P\{\sup_{\theta\in K} |\text{Log } Z_n(\theta,\tfrac{k}{n})| > \frac{\varepsilon}{12A_o}\}\right]^{-1}$$

$$\leq \frac{1}{\delta} P\{\sup_{\theta\in [0,D\sqrt{\delta}]^p} |\text{Log } Z_{[n\delta]}(\theta,1)| > \frac{\varepsilon}{12A_o}\} \times$$

$$\left[1 - \sup_{k=1,\ldots,[n\delta]} P\{\sup_{\theta\in [0,D\sqrt{\delta}]^p} |\text{Log } Z_k(\theta,1)| > \frac{\varepsilon}{12A_o}\}\right]^{-1}$$

La convergence à δ fixé de $\text{Log } Z_{[n\delta]}(\cdot,1)$ vers $\text{Log } Z(\cdot;1)$ dans $\mathcal{C}(\mathbb{R}^p)$ implique

$$P\{\sup_{\theta\in[0,D\sqrt{\delta}]^p} |\text{Log } Z_{[n\delta]}(\theta,1)| > \frac{\varepsilon}{12A_o}\} \leq 2P\{\sup_{\theta\in[0,D\sqrt{\delta}]^p} |\text{Log } Z(\theta,1)| > \frac{\varepsilon}{12A_o}\} \leq \frac{\varepsilon\delta}{4}$$

pour δ suffisamment petit $(\delta \leq \delta_o(\varepsilon,D,A_o))$ en utilisant le moment d'ordre 4 de ξ .

D'autre part le dénominateur du majorant de π_2 est supérieur à $\frac{1}{2}$ dès que :

$$\sup_n P\{\sup_{\theta\in[0,D\sqrt{\delta}]^p} |Z_n(\theta,1) - 1| > \frac{\varepsilon}{12A_o}\} \leq \frac{1}{2}$$

ce qui est réalisé en prenant δ suffisamment petit car la suite $Z_n(\cdot,1)$ est tendue .

Nous venons ainsi de montrer que :

$\forall\varepsilon > 0$, $\forall D > 0$, $\exists\delta_o(\varepsilon,D)$ et $\exists n_o(\varepsilon,D,\delta_o)$ tel que $\forall n \geq n_o$,

$$P\{\sup_{\substack{t_1,t_2\in[0,1] \\ |t_2-t_1|\leq\delta_o}} \sup_{\theta\in[0,D]^p} |Z_n(\theta,t_2) - Z_n(\theta,t_1)| > \varepsilon\} \leq \varepsilon$$

Pour les n_o premiers éléments de la suite Z_n , il est facile de trouver un δ plus petit qui convienne ; en effet, si $\delta \leq \frac{1}{n_o}$, le premier membre de l'inégalité est majoré par :

$$\sup_{n=1,\ldots,n_o} P\{\sup_{k=1,\ldots,n} \sup_{\theta\in K} |Z_n(\theta,\tfrac{k}{n}) - Z_n(\theta,\tfrac{k-1}{n})| > \frac{\varepsilon}{n\delta}\}$$

$$\leq \sup_{n=1,\ldots,n_o} P\{\sup_{k=o,\ldots,n} \sup_{\theta\in K} Z_n(\theta,\tfrac{k}{n}) > \frac{\varepsilon}{2n\delta}\}$$

$$< n_o \cdot P\{\sup_{\theta\in K} \frac{g(X,\theta_o+\theta)}{g(X,\theta_o)} > (\frac{\varepsilon}{2n_o\delta})^{\frac{q}{n_o}}\} \leq \varepsilon \quad \text{pour} \quad \delta \leq \delta'_o(\varepsilon,n_o,D) .$$

Remarque 3 : A ce stade, nous avons en particulier montré la convergence étroite à θ fixé, de la suite $Z_n(\theta,\cdot)$ dans $\mathcal{C}[0,1]$.

Pour achever cette première partie, il ne nous reste plus qu'à faire la :

Démonstration du lemme 1.2.

Pour simplifier, nous prendrons encore $K = [0,D]^p$ que nous découpons en pavés de cotés δ, d'extrémités $\underline{\theta}_j$ et $\overline{\theta}_j$, $j \in J$; on majore ainsi $P\{\sup_{t\in[0,1]} \sup_{|\theta_1-\theta_2|\leq\delta} |Z_n(\theta_2,t) - Z_n(\theta_1,t)| \geq \varepsilon\}$ par :

$$P\{\sup_{t\in[0,1]} \sup_{j\in J} \sup_{\theta\in[\underline{\theta}_j,\overline{\theta}_j]} |Z_n(\theta,t) - Z_n(\underline{\theta}_j,t)| > \frac{\varepsilon}{3}\}$$

$$\leq P\{\sup_{k=1,\ldots,n} \sup_{j\in J} Z_n(\underline{\theta}_j ; \tfrac{k}{n}) \geq A\}$$

$$+ P\{\sup_{k=1,\ldots,n} \sup_{j\in J} \sup_{\theta\in[\underline{\theta}_j,\overline{\theta}_j]} \left|\frac{Z_n(\theta ; \frac{k}{n})}{Z_n(\underline{\theta}_j ; \frac{k}{n})} - 1\right| > \frac{\varepsilon}{3A}\}$$

où $A(\varepsilon,D)$ sera fixé plus tard $(A > 1)$

$$\leq P\{\sup_{k=1,\ldots,n} \sup_{j\in J} \operatorname{Log} Z_n(\underline{\theta}_j ; \tfrac{k}{n}) \geq \operatorname{Log} A\}$$

$$+ P\{\sup_{k=1,\ldots,n} \sup_{j\in J} \sup_{\theta\in[\underline{\theta}_j,\overline{\theta}_j]} \left|\operatorname{Log} \frac{Z_n(\theta ; \frac{k}{n})}{Z_n(\underline{\theta}_j ; \frac{k}{n})}\right| > \frac{\varepsilon}{6A}\} = \pi'_1 + \pi'_2 .$$

Comme précédemment, le lemme 1.5. permet de majorer séparemment π'_1 et π'_2 :

$$\pi'_1 \leq P\{\sup_{j\in J} |\operatorname{Log} Z_n(\underline{\theta}_j ; 1)| \geq \frac{\operatorname{Log} A}{2}\} \times$$

$$\left[1 - \sup_{k=1,\ldots,n} P\{\sup_{j\in J} |\operatorname{Log} Z_n(\underline{\theta}_j ; \tfrac{k}{n})| \geq \frac{\operatorname{Log} A}{2}\}\right]^{-1}$$

le numérateur de cette dernière expression est majoré par :

$$P\{\sup_{\theta\in[0,D]^p} |\text{Log } Z_n(\theta,1)| \geqslant \frac{\text{Log } A}{2}\} \leqslant \frac{\varepsilon}{4}$$

pour $A \geqslant A_o$ et $n \geqslant n_o$ d'après (1.8).

Pour minorer le dénominateur du majorant de π_1', nous avons déjà :

$$\sup_{k=n_o,\ldots,n} P\{\sup_{\theta\in[0,D]^p} |\text{Log } Z_k(\theta\,;1)| \geqslant \frac{\text{Log } A_o}{2}\} \leqslant \frac{\varepsilon}{4}$$

les n_o premières valeurs de k utilisent la condition B1) :

$$\sup_{k=1,\ldots,n_o} P\{\sup_{\theta\in[0,D]^p} |Z_k(\theta\sqrt{\frac{n_o}{n}}\,;1) - 1| > 1 - \frac{1}{\sqrt{A_o}}\} \leqslant \frac{1}{2}$$

pour n suffisamment grand ($\frac{n_o}{n}$ petit et les mesures ν sont à marginales diffuses). Nous traitons par les mêmes arguments π_2' :

$$\pi_2' \leqslant P\{\sup_{j\in J} \sup_{\theta\in[\underline{\theta}_j,\overline{\theta}_j]} \left|\text{Log } \frac{Z_n(\theta\,;1)}{Z_n(\underline{\theta}_j\,;1)}\right| > \frac{\varepsilon}{12A_o}\} \times$$

$$\left[1 - \sup_{k=1,\ldots,n} P\{\sup_{j\in J} \sup_{\theta\in[\underline{\theta}_j,\overline{\theta}_j]} \left|\text{Log } \frac{Z_n(\theta\,;\frac{k}{n})}{Z_n(\underline{\theta}_j\,;\frac{k}{n})}\right| > \frac{\varepsilon}{12A_o}\}\right]^{-1}$$

$$\leqslant P\{\sup_{j\in J} \sup_{\theta\in[\underline{\theta}_j,\overline{\theta}_j]} \left|\frac{Z_n(\theta\,;1)}{Z_n(\underline{\theta}_j\,;1)} - 1\right| > \frac{\varepsilon}{12A_o}\} \times$$

$$\left[1 - \sup_{k=1,\ldots,n} P\{\sup_{j\in J} \sup_{\theta\in[\underline{\theta}_j,\overline{\theta}_j]} \left|\frac{Z_n(\theta\,;\frac{k}{n})}{Z_n(\underline{\theta}_j\,;\frac{k}{n})} - 1\right| > \frac{\varepsilon}{12A_o}\right]^{-1}$$

Pour utiliser les arguments précédents, il nous faut minorer en probabilité les $Z_n(\underline{\theta}_j\,;1)$ qui figurent au dénominateur :

$$P\{\sup_{j\in J} \sup_{\theta\in[\underline{\theta}_j,\overline{\theta}_j]} \left|\frac{Z_n(\theta\,;\frac{k}{n})}{Z_n(\underline{\theta}_j\,;\frac{k}{n})} - 1\right| > \frac{\varepsilon}{12A_o}\}$$

$$\leqslant P\{\inf_{j\in J} Z_n(\underline{\theta}_j\,;\frac{k}{n}) \leqslant \frac{1}{\sqrt{A_o}}\} +$$

$$+ P\{\sup_{j\in J} \sup_{\theta\in[\underline{\theta}_j,\bar{\theta}_j]} |Z_n(\theta;\tfrac{k}{n}) - Z_n(\underline{\theta}_j;\tfrac{k}{n})| \geq \frac{\varepsilon}{12A_o^{3/2}}\} = \pi_2''$$

Si k est grand, il suffit d'utiliser la convergence de $Z_k(\cdot;1)$ vers $Z(\cdot,1)$

$$\pi_2'' \leq P\{\inf_{\theta\in[0,D]^p} Z_k(\theta,1) \leq \frac{1}{\sqrt{A_o}}\} + P\{\sup_{\substack{\theta_1,\theta_2\in[0,D]^p \\ ||\theta_2-\theta_1||\leq\delta}} |Z_k(\theta_2;1)-Z_k(\theta_1;1)| > \frac{\varepsilon}{12A_o^{3/2}}\}$$

$$\leq P\{\sup_{\theta\in[0,D]^p} |\text{Log } Z_k(\theta,1)| \geq \frac{\text{Log } A_o}{2}\} + 2P\{\sup_{\substack{||\theta_2-\theta_1||\leq\delta \\ \theta_1,\theta_2\in[0,D]^p}} |Z(\theta_2;1)-(Z(\theta_1,1)| > \frac{\varepsilon}{12A_o^{3/2}}\}$$

le premier terme est majoré à l'aide de (1.8) pour $k \geq n_o(\delta_o,A_o,\varepsilon,D)$ et le second peut être rendu arbitrairement petit pour $\delta \leq \delta_o$.

Si k est inférieur à n_o , mais n grand $(\frac{n_o}{n} < \delta)$, nous avons

$$\pi_2'' \leq P\{\inf_{\theta\in[0,D\sqrt{\delta}]^p} Z_k(\theta,1) \leq \frac{1}{\sqrt{A_o}}\} + P\{\sup_{\substack{\theta_1,\theta_2\in[0,D\sqrt{\delta}]^p \\ ||\theta_2-\theta_1||\leq\delta^{3/2}}} |Z_k(\theta_2,1)-Z_k(\theta_1,1)| > \frac{\varepsilon}{12A_o^{3/2}}\}$$

$$\leq P\{\sup_{\theta\in[0,D\sqrt{\delta}]^p} |Z_k(\theta,1)-1| > \frac{\text{Log } A_o}{2}\} + P\{\sup_{\theta\in[0,D\sqrt{\delta}]^p} |Z_k(\theta,1)-1| > \frac{\varepsilon}{24A_o^{3/2}}\}$$

la condition B1) permet de rendre ces termes arbitrairement petits en prenant δ petit ; ceci achève la majoration de π_2'' donc de π_2' et par suite la démonstration du lemme .

II - Convergence dans $\mathcal{C}_\varphi$

Pour une fonction φ continue sur $[0,1]$, nulle en 0, telle que $\varphi(1) = 1$ et $\psi(t) = \frac{\varphi(t)}{\sqrt{t}}$ soit croissante sur $[0,1]$, nous allons montrer que :

$$\forall \varepsilon > 0 \ , \quad \exists M_0(\varepsilon) \text{ tel que } \forall M \geq M_0(\varepsilon) \ , \quad \forall n$$

$$P\{ \sup_{t\in]0,1]} \ \sup_{\varphi(t)||\theta||>M} Z_n(\theta,t) > \varepsilon \} \leq \varepsilon$$

Ceci nécessite des hypothèses supplémentaires sur le modèle que nous allons expliciter.

II.1. Hypothèses de régularité sur le modèle

A2. i) la fonction $\theta \rightsquigarrow g(x,\theta)$ est absolument continue sur $\overline{\Theta}$ ($\overline{\Theta}$ fermeture de Θ dans $\mathbb{R}^p$), μ-pre sque surement en x.

ii) $I(\theta) = \int \frac{\partial}{\partial\theta} g^{1/2}(x,\theta) \, {}^t\!\frac{\partial}{\partial\theta} g^{1/2}(x,\theta) d\mu(x)$ est continue sur $\overline{\Theta}$ et définie positive en θ_0.

iii) $\forall \delta > 0$, $\forall w$, $||w|| = 1$,

$$\lim_{\varepsilon\to 0} \frac{1}{\varepsilon^p} \int_{||v-\theta_0||\leq\varepsilon} d\lambda(v) \int_{\{x, |\mathrm{Log} \frac{g(x,\theta_0+\varepsilon w)}{g(x,\theta_0)}|>\delta\}} {}^t\!\frac{\partial}{\partial v} g^{1/2}(x,v) \frac{\partial}{\partial v} g^{1/2}(x,v) d\mu(x) = 0$$

où λ est la mesure de Lebesgue sur $\mathbb{R}^p$.

B2. Pour tout u inclus (au sens large) dans $\{1,2,\ldots,p\}$, il existe des constantes $\alpha_u > 1$, $\gamma_u > 0$, $a_u \geq 0$ et $B_u > 0$ telles que :

$\forall n \in \mathbb{N}$, $\forall \theta_1$, $\forall \theta_2$, $\theta_1 \leq \theta_2$, $\forall\, \theta_3$

$$E\left|Z_n(\theta_3^{\bar{u}},\ \pi_u[\theta_1,\theta_2];1)\right|^{\gamma_u} \leq B_u(1_v||\theta_1^u||_v||\theta_2^u||_v||\theta_3^{\bar{u}}||)^{a_u}\cdot\left[\lambda_u(\pi_u[\theta_1,\theta_2])\right]^{\alpha_u}$$

où λ_u est la mesure de Lebesgue sur $\mathbb{R}^{\ell(u)}$ et la notation xvy désigne le maximum des quantités x et y.

C2. Il existe $\eta > 0$ tel que $\overline{\lim_{||\theta||\to\infty}}\ ||\theta||^{\eta}\cdot\rho(\theta_o,\ \theta_o+\theta) < \infty$

où ρ désigne l'affinité de Hellinger : $\rho(\theta_o,\theta_o+\theta) = \int_{\mathcal{X}}\left[g(x,\theta_o)g(x,\theta_o+\theta)\right]^{\frac{1}{2}}d\mu(x)$

D2. Pour tout ε positif ,

$$P_{\theta_o}\left\{\sup_{\substack{\theta \\ ||\theta||>M}}\ \frac{g(X,\theta_o+\theta)}{g(X,\theta_o)} > \varepsilon\right\}$$

tend vers 0 lorsque M tend vers l'infini.

E2. La matrice Hessienne en θ de $\operatorname{Log} g(X,\theta_o+\theta)$ est p.s. continue en $\theta = 0$ et il existe $\delta > 0$ tel que $E\ \sup_{\theta,\,||\theta||\leq\delta}||\ddot{L}(X,\theta)|| < +\infty$

Remarques : i) Parmi ces hypothèses, seule B2 n' est pas exprimée en fonction de conditions sur g : nous définirons au paragraphe II.3) une condition B'2. impliquant B2. . B2. est évidemment à rapprocher de B1 .Mais la convergence étroite sur les compacts est vraie avec une mesure majorante ν quelconque : ce serait par exemple le cas si le paramétrage était tel que le modèle $g(\cdot;(\theta^{(1)},\ldots,\theta^{(p)}))$ n'est pas dérivable par

rapport à θ , mais vérifie la condition B'2 posée au II.3 par rapport au paramètre η avec $\eta^{(i)} = \zeta_i(\theta^{(i)})$, $i = 1,\ldots,p$ où les fonctions ζ_i sont croissantes sur $\mathbb{R}$ (non dérivables). En revanche, la convergence dans $\mathcal{C}_\varphi$ nécessite explicitement la mesure de Lebesgue comme mesure majorante.

ii) L'hypothèse D2 est nécessaire pour que le processus Z_1 appartienne à $\mathcal{C}_\varphi$; elle est par exemple impliquée par l'existence d'un réel M_o satisfaisant :

$$E_{\theta_o} \quad \underset{\theta,||\theta||>M_o}{\text{Sup}} \left[\frac{g(X,\theta_o+\theta)}{g(X,\theta_o)}\right]^{1/2} < \infty .$$

iii) Les hypothèses A2 sont classiques et impliquent A1 (IBRAGIMOV-KHASMINSKII [6]). En particulier A2 ii) est satisfaite dès qu'il existe r strictement supérieur à 2 tel que pour tout θ dans un voisinage de θ_o ,

$$\int || \frac{\partial}{\partial\theta} g^{\frac{1}{r}}(x,\theta)||^r \, d\mu(x) < \infty .$$

II.2. Théorème de convergence

Théorème 2 : Sous les hypothèses A2, B2, C2, D2, E2, la suite de processus de vraisemblance $Z_n(\theta,t)$ converge étroitement sous P_{θ_o} dans $\mathcal{C}_\varphi$ vers le processus $Z(\theta,t)$ dès que

$$\int_0^1 \left[\frac{\varphi(t)}{t}\right]^2 dt < \infty .$$

Remarque : La condition sur φ est à rapprocher de celle introduite pour la convergence des processus empiriques (voir par exemple [9]).

Nous allons maintenant démontrer ce théorème au moyen des lemmes suivants :

Lemme 2.1. : Sous les hypothèses A2, C2, la suite de processus $Z_n(\cdot,1)$ satisfait : $\forall N > 0$, $\exists B_N$ et $\exists n(N)$ tels que $\forall n \geqslant n(N)$, $\forall \theta$,

$$E\, Z_n(\theta,1)^{q/2} \leqslant \frac{B_N}{||\theta||^N} \quad .$$

Démonstration

En suivant IBRAGIMOV-KHASMINSKII [5] , on a d'après C2., $\forall N > 0$, $\exists U_N > 0$ et $n(N)$ tels que $\forall w$ satisfaisant $||w|| > U_N$,

$$\rho(\theta_o,\ \theta_o+w) \leqslant \frac{1}{2^N ||w||^{N\sqrt{n}}}$$

d'où $\forall \theta$, $||\theta|| > \sqrt{n}\, U_N$, $\rho^n(\theta_o,\ \theta_o + \frac{\theta}{\sqrt{n}}) \leqslant \frac{1}{||\theta||^N}$.

Etudions maintenant $||\theta||^N \cdot \rho^n(\theta_o,\ \theta_o + \frac{\theta}{\sqrt{n}})$ lorsque $||\theta|| \leqslant \sqrt{n}\, U_N$.

Le comportement à l'origine de ρ se déduit de A2. et on a :

$$\forall U > 0,\ \exists a(U) > 0 \text{ tel que } \forall w,\ ||w|| \leqslant U ,\ \rho(\theta_o,\theta_o+w) \leqslant e^{-a.||w||^2}$$

On en déduit l'inégalité : $\rho^n(\theta_o,\ \theta_o + \frac{\theta}{\sqrt{n}}) \leqslant e^{-a(U_N)\cdot||\theta||^2}$ vraie pour tout θ satisfaisant $||\theta|| \leqslant \sqrt{n}\, U_n$ et par suite l'existence d'une constante B_N telle que pour tout θ , pour tout n supérieur à $n(N)$, $\rho^n(\theta_o,\ \theta_o + \frac{\theta}{\sqrt{n}}) \leqslant \frac{B_N}{||\theta||^N}$.

L"inégalité de Hölder permet de déduire immédiatement du lemme 2.1. et de l'hypothèse B2. que : $\forall u \leqslant \{1,2,\ldots,p\}$, $\forall \gamma'_u \in]\gamma_u, \frac{q}{2}[$ et tel

que $\alpha'_u = \alpha_u \cdot \dfrac{\frac{q}{2} - \gamma'_u}{\frac{q}{2} - \gamma_u} > 1$, on a le résultat suivant :

Lemme 2.2. : Sous les hypothèses A2, B2 et C2. , on a :

$\forall N > 0$, $\exists n(N)$ et $\exists B_N(u)$ tels que $\forall n \geq n(N)$, $\forall \theta_1$, $\forall \theta_2$, $\theta_1 \leq \theta_2$, $||\theta_1||_\wedge \, ||\theta_2|| \geq 2$, $\forall \theta_3$,

$$E|Z_n(\theta_3^{\bar{u}}, \pi_u[\theta_1,\theta_2];1)|^{\gamma'_u} \leq \frac{B_N(u)}{\left[||\theta_3^{\bar{u}}||_\vee (||\theta_1^u||_\wedge ||\theta_2^u||)\right]^N} [\lambda_u(\pi_u[\theta_1,\theta_2])]^{\alpha'_u}$$

Nous considérons un réseau de $\mathbb{R}^p$: $\{\theta_I\}_{I\in\mathbb{Z}^p}$ tel que $\theta_I \leq \theta_{I+1}$, et $||\theta_{I+1} - \theta_I|| \leq 1$ avec $I+1$ désignant $I + {}^t(1,\ldots,1)$; I_* désignera celui des deux indices I et $I+1$ qui réalise $||\theta_{I_*}|| = ||\theta_I||_\vee \, ||\theta_{I+1}||$. Si la sommation ci-dessous est étendue à tous les sous-ensembles u de $\{1,2,\ldots,p\}$, y compris $u = \{1,\ldots,p\}$ et $u = \emptyset$, nous pouvons écrire :

$$P\{\sup_{\theta\in[\theta_I,\theta_{I+1}]} Z_n(\theta;1) > \varepsilon\}$$

$$\leq \sum_u P\{\sup_{\theta\in[\theta_I,\theta_{I+1}]} |Z_n(\theta_{I_*}^{\bar{u}}, \pi_u[\theta_I,\theta];1)| > \frac{\varepsilon}{2^p}\}$$

$$\leq \sum_u \left(\frac{2^p}{\varepsilon}\right)^{\gamma'_u} \frac{B_N(u).C_{\ell(u)}}{||\theta_{I_*}||^N} \quad \text{d'après les lemmes 1.4. et 2.2.}$$

$$\leq \frac{1}{\varepsilon^{\gamma_*}} \frac{B_*.C_*}{||\theta_{I_*}||^N} \quad \text{avec} \quad \gamma_* = \max_u \gamma'_u \ , \quad B_* = \max_u B_N(u) \times (2^p)^{\gamma_*+1}$$

$$\text{et} \quad C_* = \max_u C_{\ell(u)} \ .$$

La sommation sur le réseau particulier : $\theta_I = I$ conduit au :

Lemme 2.3. : Sous les hypothèses A2, B2. et C2., on a $\forall R > 2$,

$\forall N > 0$, $\exists \gamma_*$, B_N^* , C_* et $n(N)$ tels que $\forall n > n(N)$,

$$P\Big(\sup_{||\theta||>R} Z_n(\theta,1) \geqslant \varepsilon\Big) \leqslant \frac{B_N^* C_*}{\varepsilon^{\gamma_*} R^{N-p}} .$$

Nous allons partager l'ensemble :

$$\Big\{ \sup_{k=1,\ldots,n} \sup_{\varphi(\frac{k}{n})||\theta||>M} Z_n(\theta, \tfrac{k}{n}) > \varepsilon \Big\} \text{ en 2 parties que nous}$$

traiterons séparement :

$$\mathcal{D}_n(1) = \Big\{ \sup_{k=1,\ldots,n} \sup_{\substack{\varphi(\frac{k}{n})||\theta||>M \\ ||\theta|| \geqslant \sqrt{n}\,\delta}} Z_n(\theta, \tfrac{k}{n}) > \varepsilon \Big\}$$

$$\text{et } \mathcal{D}_n(2) = \Big\{ \sup_{k=1\ldots n} \sup_{\substack{\varphi(\frac{k}{n})||\theta||>M \\ ||\theta|| \leq \sqrt{n}\,\delta}} Z_n(\theta, \tfrac{k}{n}) > \varepsilon \Big\}$$

avec δ fixé par l'hypothèse E2.

La probabilité de $\mathcal{D}_n(1)$ sera contrôlée à l'aide des deux lemmes suivants :

Lemme 2.4. : Sous les hypothèses A2., B2., C2.,

$\forall \varepsilon > 0$, $\forall N > p+2$, $\exists B(N,p)$, $\exists n_o(N)$ tels que :

$\forall n \geqslant n_o(N)$, $\forall M \geqslant 2$,

$$\sum_{k=n_o(N)}^{n} P\Big(\sup_{\substack{\varphi(\frac{k}{n})||\theta||>M \\ ||\theta|| > \sqrt{n}\,\delta}} Z_n(\theta, \tfrac{k}{n}) \geqslant \varepsilon \Big) \leqslant \frac{B(N,P)}{\varepsilon^{\gamma_*} \delta^2} \left(\frac{1}{M}\right)^{N-p-2}$$

Démonstration

On rappelle que $\varphi(1) = 1$ et on distingue les deux cas :

1er cas : $M \geq \sqrt{n}\,\delta$

$$\sum_{k=k_o}^{n} P\Big\{ \sup_{\theta, ||\theta|| > \frac{M}{\varphi(\frac{k}{n})} \vee \sqrt{n}\,\delta} Z_n(\theta, \tfrac{k}{n}) \geq \varepsilon \Big\}$$

$$\leq \sum_{k=k_o}^{n} P\Big\{ \sup_{\theta, ||\theta|| > \frac{M}{\psi(\frac{k}{n})}} Z_k(\theta,1) \geq \varepsilon \Big\}$$

$$\leq \sum_{k=k_o}^{n} \frac{B_* . C_*}{\varepsilon^{\gamma_*}} \cdot \left[\frac{\psi(\frac{k}{n})}{M}\right]^{N-p} \quad \text{dès que } k_o \geq n(N) \text{ d'après le lemme 2.3.}$$

$$\leq \frac{B_* \cdot C_*}{\varepsilon^{\gamma_*}} \cdot \frac{1}{\delta^2 . M^{M-p-2}}$$

2ème cas : $M < \sqrt{n}\,\delta$

La somme précédente se décompose alors en deux parties :

$$\sum_{k=k_o}^{[n\,\varphi^{-1}(\frac{M}{\sqrt{n}\delta})]} P\Big\{ \sup_{||\theta|| > \frac{M}{\psi(\frac{k}{n})}} Z_k(\theta,1) > \varepsilon \Big\} + \sum_{k=1+[n\varphi^{-1}(\frac{M}{\sqrt{n}\delta})]}^{n} P\Big\{ \sup_{||\theta|| > \sqrt{k}\delta} Z_k(\theta,1) > \varepsilon \Big\}$$

$$\leq \sum_{k=k_o}^{[n\,\varphi^{-1}(\frac{M}{\sqrt{n}\delta})]} \frac{B_* . C_*}{\varepsilon^{\gamma_*}} \left[\frac{\psi(\frac{k}{n})}{M}\right]^{N-p} + \sum_{k=1+[n\varphi^{-1}(\frac{M}{\sqrt{n}\delta})]}^{n} \frac{B_* . C_*}{\varepsilon^{\gamma_*}} \cdot \frac{1}{(\sqrt{k}\,\delta)^{N-p}}$$

dès que $k_o \geq n(N)$ et $\sqrt{n}\,\delta > 2$.

$$\leq \frac{B_* \, C_*}{\varepsilon^{\gamma_*}} \left\{ [n\,\varphi^{-1}(\tfrac{M}{\sqrt{n}\delta})] \times \left[\frac{\psi(\varphi^{-1}(\frac{M}{\sqrt{n}\delta}))}{M}\right]^{N-p} + \frac{1}{\delta^{N-p}(\frac{N-p}{2} - 1)} \cdot \left[n\,\varphi^{-1}(\tfrac{M}{\sqrt{n}\delta})\right]^{1 - \frac{N-p}{2}} \right\}$$

$$\leq \frac{B_* \, C_*}{\varepsilon^{\gamma_*}} \cdot \frac{1}{\delta^2 \, M^{N-p-2}} \left\{ \left[\psi(\varphi^{-1}(1))\right]^{N-p-2} + \frac{1}{\frac{N-p}{2} - 1} \cdot \frac{(\frac{M}{\sqrt{n}\delta})^{N-p-2}}{\left[\varphi^{-1}(\frac{M}{\sqrt{n}\delta})\right]^{\frac{N-p-2}{2}}} \right.$$

Le résultat du lemme 2.4. se déduit maintenant du fait que $\frac{u^2}{\varphi^{-1}(u)}$ est croissante car $\psi(t) = \frac{\varphi(t)}{\sqrt{t}}$ l'est .

La majoration pour les valeurs à la fois de n et de k inférieures à $n_o(N)$ est évaluée par le :

Lemme 2.5. : $\forall \varepsilon > 0$, $\forall n_o \in \mathbb{N}$, $\forall M > 0$,

$$P\Big(\sup_{k=1,\dots,n_o \wedge n} \sup_{||\theta|| > \frac{M}{\varphi(\frac{k}{n})}} Z_n(\theta, \tfrac{k}{n}) \geq \varepsilon \Big)$$

$$\leq n_o \, P\Big(\sup_{||\theta|| > \frac{M}{\sqrt{n_o}}} \frac{g(X,\theta_o+\theta)}{g(X,\theta_o)} > \varepsilon^q \Big)$$

La démonstration est immédiate.

Pour achever la démonstration du théorème 2 , il reste à évaluer le probabilité de $\mathcal{D}_n(2)$:

Lemme 2.6. : Sous les hypothèses A2., E2.

$\forall \varepsilon \in \left]0, \frac{1}{2}\right]$, $\forall n_o \in \mathbb{N}$, $\exists M(\varepsilon) > 0$ et $B > 0$ tels que

$$\sup_{n \geq n_o} P\Big(\sup_{n_o \leq k \leq n} \sup_{\substack{||\theta|| < \delta\sqrt{n} \\ ||\theta|| > \frac{M}{\varphi(\frac{k}{n})}}} Z_n(\theta, \tfrac{k}{n}) > \varepsilon \Big) \leq \varepsilon + \frac{B}{M^2} \qquad (2.1)$$

Démonstration

$$\sup_{n \geq n_o} P\Big\{ \sup_{n_o \leq k \leq n} \sup_{\frac{M}{\varphi(\frac{k}{n})} < ||\theta|| < \delta\sqrt{n}} Z_n(\theta, \tfrac{k}{n}) > \varepsilon \Big\}$$

$$= \sup_{n \geq n_o} P\Big\{ \sup_{n_o \leq k \leq n} \sup_{\frac{M}{\psi(\frac{k}{n})} < ||\theta|| < \delta\sqrt{k}} \sum_{i=1}^{k} \operatorname{Log} \frac{g(X_i, \theta_o + \frac{\theta}{\sqrt{k}})}{g(X_i, \theta_o)} > q \operatorname{Log} \varepsilon \Big\}$$

$$= \operatorname{Sup} P\{ \operatorname*{Sup}_{n \geq \frac{M^2}{\delta^2}} \; \operatorname*{Sup}_{n\, \varphi^{-1}(\frac{M}{\sqrt{n}\delta}) \leq k \leq n} \; \frac{M}{\psi(\frac{k}{n})} < ||\theta|| < \delta\sqrt{k} \; \frac{t_\theta}{\sqrt{k}} \sum_{i=1}^{k} \dot{L}(X_i, 0) +$$

$$\frac{1}{2k}\, t_\theta \sum_{i=1}^{k} \ddot{L}(X_i, \frac{\theta_k^*}{\sqrt{k}})\theta \geq q \operatorname{Log} \varepsilon\} \qquad (2.2)$$

avec $\dot{L}(X,\theta)$ désignant le gradient en θ de $\operatorname{Log} g(X,\theta_o+\theta)$ et $\theta_k^* \in [0,\theta]$.

Il importe maintenant de contrôler le terme quadratique. Pour cela introduisons $\bar{\lambda}(A)$ et $\underline{\lambda}(A)$ les, respectivement, plus grande et plus petite valeurs propres de la matrice A. L'hypothèse E2 implique

$$- E_{\theta_o} \operatorname*{Sup}_{||\theta|| \leq \delta} \bar{\lambda}(\ddot{L}(X,\theta)) - (1 - \frac{\varepsilon}{2})\underline{\lambda}(I(\theta_o)) > 0 \qquad (2.3)$$

D'autre part, la loi forte des grands nombres implique l'existence d'un entier $n_o(\varepsilon)$ tel que :

$$P\{\forall k \geq n_o(\varepsilon),\; E_{\theta_o} \operatorname*{Sup}_{||\theta|| < \delta_o} \bar{\lambda}(\ddot{L}(X,\theta)) - \frac{1}{k} \sum_{i=1}^{k} \operatorname*{Sup}_{||\theta|| < \delta_o} \bar{\lambda}(\ddot{L}(X_i,\theta)) + \frac{\varepsilon}{2}\, \underline{\lambda}(I(\theta_o)) > 0\}$$

$$\geq 1 - \varepsilon \qquad (2.4)$$

Les assertions (2.3) et (2.4) permettent de majorer (2.2) dès que $n\, \varphi^{-1}(\frac{M}{\sqrt{n}\delta}) \geq n_o(\varepsilon)$ par :

$$\varepsilon + \operatorname{Sup} P\{ \operatorname*{Sup}_{n > \frac{M^2}{\delta^2}} \; \operatorname*{Sup}_{n\, \varphi^{-1}(\frac{M}{\sqrt{n}\delta}) \leq k \leq n} \; \frac{M}{\psi(\frac{k}{n})} \leq ||\theta|| \leq \delta\sqrt{k} \; \frac{t_\theta}{\sqrt{k}} \sum_{i=1}^{k} \dot{L}(X_i, 0) - \alpha \frac{1-\varepsilon}{2} t_\theta I(\theta_o)\theta$$

$$> q \operatorname{Log} \varepsilon\} \qquad (2.5)$$

avec $\alpha = \dfrac{\underline{\lambda}(I(\theta_o))}{\bar{\lambda}(I(\theta_o))}$.

A k fixé, le suprémum (absolu) en θ est atteint au point :

$$\tilde{\theta}_k = \frac{1}{\alpha(1-\varepsilon)}\, I(\theta_o)^{-1}\, \frac{1}{\sqrt{k}} \sum_{i=1}^{k} \dot{L}(X_i, 0)$$

Et la majoration de (2.5) se poursuit par :

$$\varepsilon + \sup_{n \geq \frac{M^2}{\delta^2}} P\{ \bigcup_{n\,\varphi^{-1}(\frac{M}{\sqrt{n}\delta}) \leq k \leq n} ||\tilde{\theta}_k|| > \frac{M}{\psi(\frac{k}{n})} \}$$

$$+ \sup_{n \geq \frac{M^2}{\delta^2}} P\{ \sup_{n\,\varphi^{-1}(\frac{M}{\sqrt{n}\delta}) \leq k \leq n} \sup_{||\theta|| = \frac{M}{\psi(\frac{k}{n})}} \frac{t_\theta}{\sqrt{k}} \sum_{i=1}^{k} \dot{L}(X_i,0) - \frac{1-\varepsilon}{2}\alpha\, {}^t\theta\, I(\theta_o)\theta > q \operatorname{Log} \varepsilon \}$$

en décomposant le second terme suivant les coordonnées, le majorant devient :

$$\varepsilon + \sup_{n \geq \frac{M^2}{\delta^2}} \sum_{\ell=1}^{p} P\{ \sup_k \frac{\psi(\frac{k}{n}).||I(\theta_o)^{-1}||}{M.\sqrt{k}(1-\varepsilon)\alpha} \Big| \sum_{i=1}^{k} \dot{L}^{(\ell)}(X_i,0) \Big| > 1 \}$$

$$+ \sup_{n \geq \frac{M^2}{\delta^2}} P\{ \sup_k \frac{M}{\psi(\frac{k}{n})} \Big|\Big| \frac{1}{\sqrt{k}} \sum_{i=1}^{k} \dot{L}(X_i,0) \Big|\Big| \geq \operatorname{Log} \varepsilon^q + \frac{1-\varepsilon}{2}\alpha \cdot \frac{M^2}{||I(\theta_o)^{-1}|| \left[\psi(\frac{k}{n})\right]^2} \} \quad (2.6)$$

L'inégalité de BIRNBAUM-MARSHALL [3] sur les martingales majore (2.6) par :

$$\varepsilon + \sup_{n \geq \frac{M^2}{\delta^2}} \sum_{\ell=1}^{p} \sum_{k=\left[n\varphi^{-1}(\frac{M}{\delta\sqrt{n}})\right]}^{n} \frac{||I(\theta_o)^{-1}||^2 \cdot ||I(\theta_o)||}{\alpha^2(1-\varepsilon)^2 M^2 n} \left[\frac{\varphi(\frac{k}{n})}{\frac{k}{n}}\right]^2$$

$$+ \sup_{n \geq \frac{M^2}{\delta^2}} \frac{p||I(\theta_o)||M^2}{n} \sum_{k=\left[n\varphi^{-1}(\frac{M}{\sqrt{n}\delta})\right]}^{n} \varphi(\frac{k}{n})^{-2} \left[\frac{1-\varepsilon}{4} \frac{M^2}{||I(\theta_o)^{-1}|| \left[\psi(\frac{k}{n})\right]^2}\right]^{-2}$$

pour M assez grand et

$$\leq \varepsilon + \frac{p}{\alpha^2(1-\varepsilon)^2} ||I(\theta_o)^{-1}||^2 \, ||I(\theta_o)|| \cdot \int_o^1 \left[\frac{\varphi(t)}{t}\right]^2 dt \cdot \frac{1}{M^2}$$

II.3. Dans les hypothèses du théorème 2, figure l'hypothèse B2 qui n'est pas directement exprimée en fonction de conditions sur la densité g ; nous allons donc consacrer ce paragraphe à exhiber des hypothèses sur g susceptibles d'impliquer B2.

Pour un sous-ensemble u non vide de $\{1,2,\ldots,p\}$, on notera Δ_u ou $\frac{\partial^{\ell(u)}}{\partial\theta^u}$ l'opérateur de dérivation par rapport aux coordonnées d'indices appartenant à u :

$$\text{si} \qquad u = \{j_1, j_2, \ldots, j_{\ell(u)}\} \ , \ \Delta_u = \frac{\partial^{\ell(u)}}{\partial\theta^{(j_1)} \ldots \partial\theta^{(j_{\ell(u)})}}$$

Nous pouvons maintenant expliciter la condition :

B'2) Il existe $a' > 0$, un entier $m \geq 1$ et un entier pair q tels que $m.q > p$ et $\sup_{\theta\in\Theta} \frac{1}{1+||\theta||^{a'}} \int |\Delta_u \, g^{\frac{1}{q}}(x,\theta_o+\theta)|^q d\mu(x) < \infty$

pour tout sous ensemble u vérifiant $\ell(u) \leq m$.

Remarque : dans le cas particulier où $m = 1$, la condition q entier pair peut-être facilement supprimée, il reste $q > p$ et IBRAGIMOV-KHASMINSKII montre dans [6] , que cette dernière hypothèse implique l'existence d'une constante B telle que :

$$\forall R > 0 \ , \quad \sup_{\substack{\theta_1,\theta_2 \\ ||\theta_1||\leq R, ||\theta_2||\leq R}} E|Z_n(\theta_2,I)-Z_n(\theta_1,I)|^q \leq B(1+R^a)||\theta_2-\theta_1||^q \quad \text{avec} \quad q > p$$

il est alors facile de voir que cela implique que la condition B2 est satisfaite pour tout sous ensemble u de $\{1,2,\ldots,p\}$ avec $a_u = a$,

$\gamma_u = \gamma = q$, $\alpha = \frac{q}{p}$ et $\alpha_u = \frac{q}{\ell(u)}$: il suffit de faire un découpage des pavés en cubes pour s'en convaincre.

Plus généralement, montrons le :

Lemme 2.7. : Si la densité $g(\cdot,\theta)$ satisfait l'hypothèse B'2), alors la suite de processus $Z_n(\cdot\,;1)$ satisfait la condition B2 avec :

$$a_u = a = a'.m, \quad \gamma_u = \gamma = q, \quad \alpha = \frac{m.q}{p} \quad \text{et} \quad \alpha_u = \frac{m.q}{m_v \ell(u)}$$

Démonstration

Pour montrer que $\forall u \subseteq \{1,\dots,p\}$,

$$E\left[Z_n(\theta_3^{\bar{u}},\pi_u[\theta_1,\theta_2];1)\right]^q \leq B_u\left[1_v\|\theta_3^{\bar{u}}\|_v \|\theta_1^u\|_v \|\theta_2^u\|\right]^a.\left[\lambda_u(\pi_u[\theta_1,\theta_2])\right]^{\alpha_u} ,$$

nous sommes conduits à considérer les sous-ensembles u' inclus dans u et vérifiant $\ell(u') \leq m$; nous pouvons écrire :

$$Z_n(\theta_3^{\bar{u}},\theta_4^{u\setminus u'},\pi_u[\theta_1,\theta_2];1) = \int\dots\int_{\pi_u[\theta_1,\theta_2]} \Delta_u Z_n((\theta_3^{\bar{u}},\theta_4^{u\setminus u'},\theta^{u'});1)d\lambda_u(\theta^{u'})$$

L'inégalité de Holder fournit une première majoration :

$$E\left[Z_n(\theta_3^{\bar{u}},\theta_4^{u\setminus u'},\pi_u[\theta_1,\theta_2];1\right]^q \leq E\left[\lambda_u(\pi_u[\theta_1,\theta_2])\right]^{q-1}$$

$$\int\dots\int_{\pi_u[\theta_1,\theta_2]}\left[\Delta_u Z_n((\theta_3^{\bar{u}},\theta_4^{u\setminus u'},\theta^{u'});1)\right]^q d\lambda_u(\theta^{u'})$$

$$\leq \left[\lambda_u(\pi_u[\theta_1,\theta_2])\right]^q . \sup_{\theta^{u'}\in\pi_u[\theta_1,\theta_2]} E\left[\Delta_u Z_n((\theta_3^{\bar{u}},\theta_4^{u\setminus u'},\theta^{u'});1)\right]^q$$

Si $\ell(u)$ est inférieur ou égal à m, on prend seulement $u' = u$; et si

$\ell(u)$ est supérieur à m, on prend tous les sous-ensembles u' inclus dans u et vérifiant $\ell(u') = m$; on remarque que $Z_n(\theta_3^{\bar{u}}, \pi_u[\theta_1,\theta_2] ; 1)$ peut s'écrire comme la somme de $2^{\ell(u)-m}$ termes de la forme $Z_n(\theta_3^{\bar{u}}, \theta_1^{u \setminus u'}, \pi_{u'}[\theta_1,\theta_2] ; 1)$ ce qui conduit alors à la majoration :

$$E\left[Z_n(\theta_3^{\bar{u}}, \pi_u[\theta_1,\theta_2] ; 1)\right]^q \leq \left[\lambda_u(\pi_u[\theta_1,\theta_2])\right]^{\frac{m.q}{\ell(u)}} \cdot 2^{q[\ell(u)-m]} \times \ldots$$

$$\sup_{\substack{u' \subset u \\ \ell(u')=m}} \sup_{\theta^u \in \pi_u[\theta_1,\theta_2]} E\left[\Delta_{u'} Z_n((\theta_3^{\bar{u}}, \theta^u) ; 1)\right]^q$$

Pour achever la preuve du lemme 2.7., il ne reste plus qu'à montrer le :

Lemme 2.8. : L'hypothèse B'2) avec q pair implique :

$$\sup_{\substack{u' \subseteq u \\ \ell(u') \leq m}} \sup_{\theta^u \in \pi_u[\theta_1,\theta_2]} E\left[\Delta_{u'} Z_n((\theta_3^{\bar{u}}, \theta^u) ; 1)\right]^q \leq B_u(1_v \|\theta_3^{\bar{u}}\|_v \|\theta_1^u\|_v \|\theta_2^u\|)^a$$

Démonstration

Il suffit par raison de symétrie sur les coordonnées d' étudier $u' = \{1,2,\ldots,m'\}$ avec $m' \leq m$. On peut écrire :

$$\frac{\partial}{\partial \theta^{(1)}} Z_n(\theta,1) = \Delta_1 \left\{ \prod_{k=1}^{n} \frac{\frac{1}{g^q}(X_k, \theta_o + \frac{\theta}{\sqrt{n}})}{\frac{1}{g^q}(X_k, \theta_o)} \right\}$$

$$= \sum_{i_1=1}^{n} \prod_{k=1}^{n} \Delta_1^{\delta(i_1,k)} \frac{\frac{1}{g^q}(X_k, \theta_o + \frac{\theta}{\sqrt{n}})}{\frac{1}{g^q}(X_k, \theta_o)}$$

où $\delta(\cdot,\cdot)$ désigne le symbole de Kronecker

$$\Delta_{u'} Z_n(\theta,1) = \sum_{i_1=1}^{n} \ldots \sum_{i_{m'}=1}^{n} \prod_{k=1}^{n} \Delta_1^{\delta(i_1,k)} \ldots \Delta_{m'}^{\delta(i_{m'},k)} \left[\frac{\frac{1}{g^q}(X_k, \theta_o + \frac{\theta}{\sqrt{n}})}{\frac{1}{g^q}(X_k, \theta_o)} \right]$$

En élevant à la puissance q, on obtient le développement suivant :

$$\left[\Delta_u, Z_n(\theta,1)\right]^q = \sum_{i_1^1=1}^{n} \cdots \sum_{i_{m'}^1=1}^{n} \sum_{i_1^2=1}^{n} \cdots \sum_{i_{m'}^q=1}^{n} \prod_{k=1}^{n} \prod_{\ell=1}^{q} \Delta_1^{\delta(i_1^\ell,k)} \cdots \Delta_{m'}^{\delta(i_{m'}^\ell,k)} \left[\frac{g^{\frac{1}{q}}(X_k,\theta_o+\frac{\theta}{\sqrt{n}})}{g^{\frac{1}{q}}(X_k,\theta_o)}\right]$$

et par intégration :

$$E\left[\Delta_u, Z_n(\theta,1)\right]^q = \frac{1}{n^{\frac{m'.q}{2}}} \cdot \sum_{i_1^1=1}^{n} \cdots \sum_{i_{m'}^q=1}^{n} \prod_{k=1}^{n} \int \prod_{\ell=1}^{q} \Delta_1^{\delta(i_1^\ell,k)} \cdots \Delta_{m'}^{\delta(i_{m'}^\ell,k)} \; g^{\frac{1}{q}}(x_k,\theta_o+\frac{\theta}{\sqrt{n}})\, d\mu(x_k)$$

nous avons une somme de $n^{m'.q}$ termes dont beaucoup sont nuls car pour tout $j = 1,\ldots,p$,

$$\int \Delta_j^1 \; g^{\frac{1}{q}}(x,\theta_o+\frac{\theta}{\sqrt{n}}) \cdot g^{\frac{q-1}{q}}(x,\theta_o+\frac{\theta}{\sqrt{n}})\, d\mu(x) = \frac{1}{q} \int \frac{\partial}{\partial\theta^{(j)}} g(x,\theta_o+\frac{\theta}{\sqrt{n}})\, d\mu(x) = 0$$

Par conséquent, ne seront éventuellement non nuls que les termes d'indices $(i_1^1,\ldots,i_{m'}^1,\ldots,i_1^q,\ldots,i_{m'}^q)$ vérifiant :

$$\forall k = 1,\ldots,n \;, \quad \sum_{\ell=1}^{q} \sum_{j=1}^{m'} \delta(i_j^\ell,k) \neq 1$$

il faut donc que les $m'.q$ dérivations soient groupées par "paquets de taille au moins 2" sur les indices k ; le nombre de paquets étant inférieur ou égal à $\frac{m'.q}{2}$, le nombre de termes éventuellement non nuls est inférieur à $C(m',q).n^{\frac{m'.q}{2}}$ et chacun des termes restant est majoré facilement :

. pour les indices k tels que $\sum_{\ell=1}^{q} \sum_{j=1}^{m'} \delta(i_j^\ell,k) = 0$, l'intégrale vaut 1

. pour les indices k tels que $\sum_{\ell=1}^{q} \sum_{j=1}^{m'} \delta(i_j^\ell,k) \geqslant 2$, l'inégalité de Holder donne :

$$\left| \int \prod_{\ell=1}^{q} \Delta_1^{\delta(i_1^\ell,k)} \cdots \Delta_{m'}^{\delta(i_{m'}^\ell,k)} \frac{1}{g^q}(x_k, \theta_o + \frac{\theta}{\sqrt{n}}) d\mu(x_k) \right|$$

$$\leq \prod_{\ell=1}^{q} \left[\int \left| \Delta_1^{\delta(i_1^\ell,k)} \cdots \Delta_{m'}^{\delta(i_{m'}^\ell,k)} \frac{1}{g^q}(x, \theta_o + \frac{\theta}{\sqrt{n}}) \right|^q d\mu(x) \right]^{\frac{1}{q}}$$

et le lemme en découle immédiatement.

BIBLIOGRAPHIE

[1] P.J. BICKEL - M.J. WICHURA : "Convergence criteria for multiparameter stochastic processus and some applications" A.M.S. n° 42 (1971) p. 1656-1670.

[2] P. BILLINGSLEY : "Convergence of probability measures" - Wiley.

[3] Z.W. BIRNBAUM - A.W. MARSHALL : "Some multivariate Chebyshev inequalities with extensions to continuous parameter processes" A.M.S. n° 32 (1961) p. 687-703.

[4] N.N. CHENSOV : "Limit theorems for some classes of random functions" Proceedings All Union Conf. Theory Proba and Math. Statistics (Erevan 1958). Izdat. Akad. Nauk Armjan SSR. Erevan. 1960 p. 280-285.

[5] I.A. IBRAGIMOV - R.Z. KHASMINSKII : "Asymptotic behaviour of statistical estimators in the smooth case I. study of the likelihood ratio" Theory of probability and applications n° 17 (1972) p. 445-462.

[6] I.A. IBRAGIMOV - R.Z. KHASMINSKII : "Statistical Estimation - Asymptotic theory" Springer-Verlag - App lications of mathematics n° 16.

[7] L. LECAM : "On the assumption used to prove asymptotic normality of maximum likelihood estimates" A.S.M. n° 41 (1970) p. 802-828.

[8] T. PARTHASARATHY : "Selection theorems and their applications" Lecture Notes in Mathematics - Springer-Verlag n° 263.

[9] R. PYKE - G.R. SHORACK : "Weak convergence of a two-sample empirical process and a new approach to Chernoff-Savage theorems" A.M.S. n° 39 (1968).

[10] M. YOR : "Le drap brownien comme limite en loi de temps locaux linéaires" (1981) à paraître.

[11] J. DESHAYES - D. PICARD : "Testing for a change point in statistical models". Prépublication d'Orsay (1980) t. 52.

[12] D. PICARD - J. DESHAYES : "Rupture dans les modèles de régression : loi asymptotique des tests et estimateurs du maximum de vraisemblance". Prépublication d'Orsay (1982).

RUPTURES DANS LES MODELES DE REGRESSION

LOIS ASYMPTOTIQUES DES TESTS ET ESTIMATEURS DU MAXIMUM DE VRAISEMBLANCE

Jean DESHAYES, Dominique PICARD.

Le but de cet article est d'étudier le comportement asymptotique des méthodes de vraisemblance dans les modèles de régression avec paramètres auxiliaires, perturbés par une rupture au cours de l'expérience. A l'aide d'un théorème d'invariance sur le processus de vraisemblance sous des conditions relativement faibles sur le modèle, nous étudierons de façon systématique le comportement des tests de vraisemblance (pour tester la présence d'une rupture) et les modifications qu'il convient d'y apporter ainsi que les lois asymptotiques des estimateurs du maximum de vraisemblance des paramètres qui caractérisent une rupture (l'instant d'apparition et l'amplitude).

The aim of this paper is to investigate the asymptotic behaviour of likelihood method's in the case of regression models with auxillary parameters, when a change point may occur during the experience. Around an invariance theorem on likelihood process under reasonnably weak assumptions on the model, we systematically investigate the asymptotic behaviour of likelihood tests (for testing whether there is a change point or not) and helpful modifications which can be given in this case, the asymptotic distributions of maximum likelihood estimators of the parameter characterizing a change (time of change and amplitude).

INTRODUCTION

Toute étude de comportement asymptotique (test ou estimation) dans les modèles de rupture peut être envisagée essentiellement sous deux points de vue correspondant à deux techniques probabilistes différentes :

Le premier point de vue consiste à laisser au problème sa discrétisation initiale par rapport au temps (pour peu évidemment qu'il y ait discrétisation initiale) et à obtenir des lois asymptotiques sur l'instant de rupture directement sur $\mathbb{Z}$ ou $\mathbb{N}$. Ce point de vue a déjà été étudié sur différents exemples par des méthodes de fluctuations, citons principalement Hinkley [7], [8].

Le second point de vue consiste à faire converger la suite d'expériences vers un modèle en temps continu et les résultats s'obtiennent toujours en rajoutant le paramètre temps aux paramètres usuels au moyen d'un principe d'Invariance : Citons pour exemples : Bhattacharya [2], [3], Bhattacharya-Johnson [4], Hawkins [6], Pettitt [13], Sen [16], ...

Dans ce point de vue intervient toujours une difficulté qui tient essentiellement à la possibilité de rupture intervenant au bord (début ou fin) de l'intervalle d'observation et du lien qui devient alors fondamental entre les paramètres "instant de rupture" et "amplitude" du changement. Il convient alors d'employer des techniques généralement fines, proches de la loi du logarithme itéré pour surmonter cette difficulté.

Notre but ici était d'étudier le plus profondément possible ces types de difficultés, leurs conséquences et leurs remèdes tant sur le plan de l'estimation que sur celui des tests. C'est pourquoi nous avons choisi d'étudier les phénomènes de rupture dans le cadre des méthodes de

vraisemblance (là, où l'estimation des paramètres se pose dans sa plus grande technicité) sur des modèles de régression par rapport à des fonctions dépendant du temps, ce qui en compliquant un peu le problème, l'explique plus précisément.

Il est alors intéressant de constater que la présence d'une rupture ne change pas fondamentalement le problème statistique, en ce sens que sous des conditions sur le modèle très peu différentes des conditions standards, (Les conditions que l'on impose sont à peu de choses près celles d'Ibraginov-Khashminskii dans [9], [10]) des renormalisations appropriées permettent à la fois de faire des tests de rapport de vraisemblance et d'obtenir la loi limite des estimateurs du maximum de vraisemblance. Par ailleurs, il ne fait aucun doute que ces techniques employées sur les méthodes de vraisemblance peuvent s'étendre à beaucoup de cas standards (Certains types de non régularité en variables indépendantes équidistribuées, séries chronologiques, etc ...).

La première partie de cet article est consacrée à l'obtention d'un théorème d'Invariance sur le processus de vraisemblance dans le cadre d'une régression. Le théorème est, en réalité une conséquence du théorème 2 obtenu dans [21], bien que ce dernier en apparaisse un cas particulier. Comme pour celui-ci, le principe d'Invariance est faux dans $\mathscr{C}_o$ sous les conditions les plus générales et il convient de le montrer dans un espace élargi : $\mathscr{C}_\Phi$ établissant une pondération entre les paramètres t et θ à l'infini. Ici, cependant la fonction Φ est choisie multidimensionnelle de façon à adapter plus précisément à la régression la normalisation au bord. Ce théorème est le pivot de ce travail. Nous reportons sa démonstration dans la quatrième partie parce qu'elle est assez longue et que, nous préférons énoncer d'abord les résultats statis-

tiques. Néanmoins, les techniques employées dans cette démonstration se retrouveront largement au cours des paragraphes suivants et certaines démonstrations y feront directement référence.

Le second paragraphe est consacré au problème de tester l'existence d'une rupture : une conséquence immédiate du théorème d'Invariance est que la statistique du rapport de vraisemblance, au lieu d'être asymptotiquement un X^2 (comme c'est le cas par exemple si l'on teste l'existence d'une rupture à un instant $k_o(n) = nt_o$ connu), tend vers l'infini à cause des problèmes de bords. La seule manière raisonnable de remédier à cet inconvénient est de renormaliser la statistique de vraisemblance par une fonction qui pénalise les bords. Le degré de finesse de cette pondération dépendra des possibilités du modèle.

Enfin, la troisième partie est consacrée à l'estimation globale par le maximum de vraisemblance des paramètres du changement (instant et amplitude).

I.- MODELE - EXEMPLES - THEOREME D'INVARIANCE

Soit f une fonction matricielle réelle de dimension $r \times p_1$ continue sur $[0,1]$. Nous poserons :

$$F_n(k) = \frac{1}{n} \sum_{i=1}^{k} f(\tfrac{i}{n})' f(\tfrac{i}{n}) , \quad k = 1,\ldots,n$$

et $$F(t) = \int_0^t f(u)' f(u)du , \quad t \in [0,1] ;$$

(M' désignant la matrice transposée de M)

Soit Σ un ouvert de $\mathbb{R}^{p_2}$, N un ouvert de $\mathbb{R}^{p_1}$, $\Theta = N \times \Sigma \subset \mathbb{R}^p$, $p=p_1+p_2$.

Soit $\{P_\sigma, \sigma \in \Sigma\}$ une famille de lois de probabilité sur $\mathcal{X} = (\mathbb{R}^r, \mathcal{B}(\mathbb{R}^r))$ absolument continues par rapport à la mesure de Lebesgue λ_r de densité $g(\cdot,\sigma)$. Pour θ dans Θ , $\theta' = (\eta', \sigma')$ nous définissons l'hypothèse $H_{o,n}(\theta)$ (: Régression linéaire sans rupture) par l'observation des vecteurs aléatoires de $\mathbb{R}^r$:

$X_1,\ldots,X_n$ indépendants et tels que :

$X_i - f(\frac{i}{n})\eta$ suit la loi P_σ pour tout i, $1 \leqslant i \leqslant n$

Nous appelerons rupture dans le modèle l'existence d'un instant k, $1 \leqslant k < n$ et de vecteurs θ^*, θ^{**} de Θ($\theta^{*\prime} = (\eta^{*\prime}, \sigma^{*\prime})$, $\theta^{**\prime} = (\eta^{**\prime}, \sigma^{**\prime})$, $\theta^* \neq \theta^{**}$} (Hypothèse $H_{1n}(\frac{k}{n}, \theta^*, \theta^{**})$) l'observation des vecteurs aléatoires de $\mathbb{R}^r$, $X_1,\ldots,X_n$ indépendants et tels que :

$X_i - f(\frac{i}{n})\eta^*$ suit la loi P_{σ^*} pour $1 \leqslant i \leqslant k$

$X_i - f(\frac{i}{n})\eta^{**}$ suit la loi $P_{\sigma^{**}}$ pour $k < i \leqslant n$.

Pour simplifier, nous noterons souvent $P_{H_{o,n}(\theta)}$ (resp. $P_{H_{1,n}(\frac{k}{n},\theta^*,\theta^{**})}$) la loi de $X_1,\ldots,X_n$ sous l'hypothèse $H_{o,n}(\theta)$

(resp. $H_{1,n}(\frac{k}{n}, \theta^*, \theta^{**})$). Le modèle ci-dessus contient en particulier les trois exemples suivants que nous détaillerons dans la suite :

Exemple 1 : $p_1 = 0$; il s'agit de déceler ou d'estimer les paramètres d'une rupture dans un modèle statistique classique multidimensionnel. Dans [21] , nous avons établi un théorème d'invariance précisément pour le processus de vraisemblance de ce modèle.

Exemple 2 : $p_2 = 0$; il s'agit de déceler ou d'estimer les paramètres de rupture dans un modèle de régression sans paramètre auxiliaire (par exemple, variance de la loi connue).

Exemple 2bis : modèle de translation-homothétie : $r = 1$ et $p_2 = 1$ avec :

$$X_i \text{ de densité } \frac{1}{\sigma} g\left(\frac{\cdot - f(\frac{i}{n})\eta}{\sigma}\right) .$$

Exemple 3 : $r = 1$ et p_2 quelconque : ce modèle conduira à des traitements plus simples du fait que les X_i sont unidimensionnels.

Nous introduisons naturellement le processus de vraisemblance $Z_n^f(\theta ; t)$ de $\mathscr{C}\left(\mathbb{R}^p \times [0,1]\right)$ défini comme suit :

pour tout θ appartenant à $\mathbb{R}^p$, $Z_n^f(\theta,\cdot)$ est la ligne polygonale joignant les points $(0;1)$ et $\left(\frac{k}{n}, \left[\prod_{i=1}^{k} \frac{P(X_i ; \theta_o + \frac{\theta}{\sqrt{n}})}{P(X_i ; \theta_o)}\right]^{\frac{1}{q}}\right)$, $k = 1,\ldots,n$, avec $P(X_i,\theta)=g(X_i - f(\frac{i}{n})\eta;\sigma)$.

Le point $\theta_o' = (\eta_o', \sigma_o')$ désigne la vraie valeur du paramètre et le réel positif q sera précisé plus tard par les hypothèses sur la densité $g(\cdot,\theta)$. Nous verrons que les lois limites des statistiques de tests et des estimateurs sont toutes déduites du théorème 1 qui établit la conver-

gence de la suite des processus de vraisemblance Z_n^f. Introduisons donc le cadre et les hypothèses qui sont nécessaires.

Soit ϕ une fonction matricielle symétrique de $[0,1]$ dans $\mathbb{R}^{p\times p}$ qui est continue, nulle en 0, définie positive sur $]0,1]$ et croissante (la relation d'ordre utilisée dans tout l'article sera celle des matrices symétriques définies positives : ϕ, ϕ_1 symétriques ; $\phi > \phi_1 \iff \phi - \phi_1$ définie positive) ; généralisant la démarche de [21], nous appelerons $\mathscr{C}_\Phi$ l'ensemble des fonctions z continues sur $\mathbb{R}^p \times [0,1]$ telles que :

$$z(\theta,0) \equiv 1$$

$$\lim_{M\to\infty} \sup_{t\in]0,1[} \sup_{\theta,\ {}^t\theta\phi(t)\theta\geq M} |z(\theta,t)| = 0$$

Cet ensemble $\mathscr{C}_\Phi$ est muni naturellement de la topologie de la convergence uniforme sur les compacts avec "équiannulation à l'infini" : on dit qu'une suite z_n converge dans $\mathscr{C}_\Phi$ vers z si et seulement si :

i) pour tout ε positif, pour tout compact K de $\mathbb{R}^p$, il existe $n_o(\varepsilon,K)$ tel que pour tout n supérieur à $n_o(\varepsilon,K)$,

$$\sup_{(\theta,t)\in K\times[0,1]} |z_n(\theta,t) - z(\theta,t)| \leq \varepsilon \quad .$$

ii) pour tout ε positif, il existe $M_o(\varepsilon)$ tel que pour tout M supérieur à $M_o(\varepsilon)$, pour tout n,

$$\sup_{(\theta,t),\ {}^t\theta\phi(t)\theta\geq M} |z_n(\theta,t)| \leq \varepsilon$$

Les hypothèses portant sur la densité g sont traduites dans les cinq conditions A à E et l'hypothèse sur le regresseur dans la condition F -

A : i) la fonction $\sqrt{g}(\cdot,\cdot)$ est absolument continue sur $\mathbb{R}^r \times \Sigma$ et prolongée par continuité sur $\mathbb{R}^r \times \overline{\Sigma}$ où $\overline{\Sigma}$ dés igne la fermeture de Σ dans $\mathbb{R}^{p_2}$,

ii) l'information de Fisher existe et

$$\mathcal{I}(\sigma) = \int 4 \begin{pmatrix} \frac{\partial}{\partial x} g^{1/2}(x,\sigma) \\ \\ \frac{\partial}{\partial \sigma} g^{1/2}(x,\sigma) \end{pmatrix} \begin{pmatrix} \frac{\partial}{\partial x} g^{1/2}(x,\sigma) \\ \\ \frac{\partial}{\partial \sigma} g^{1/2}(x,\sigma) \end{pmatrix}' d\lambda_r(x)$$

est une matrice définie positive sur Σ et continue sur $\overline{\Sigma}$.

iii) $\forall \delta > 0$, $\forall w = (w_1,w_2) \in \mathbb{R}^{r+p_2}$, $\|w\| = 1$,

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon^p} \int_{\|v_1\|<\varepsilon} d\lambda_r(v_1) \int_{\|v_2-\sigma_o\|<\varepsilon} d\lambda_{p_2}(v_2)$$

$$\int_{\{x,\,|\log \frac{g(x+\varepsilon w_1,\sigma_o+\varepsilon w_2)}{g(x,\sigma_o)}| > \delta\}} \frac{\partial}{\partial v} g^{\frac{1}{2}}(x-v_1,v_2)' \frac{\partial}{\partial v} g^{\frac{1}{2}}(x-v_1,v_2) d\lambda_r(x) = 0$$

avec $v' = (v_1', v_2')$.

B : Pour une combinaison u_1 avec répétition de $\{1,2,\ldots,p_1\}$ de taille $\ell(u_1)$ et une combinaison u_2 sans répétition de $\{1,2,\ldots,p_2\}$ de taille $\ell(u_2)$ on note $\Delta_u = \Delta_{(u_1,u_2)}$ l'opérateur de dérivation par rapport aux coordonnées de x d'indices appartenant à u_1 et aux coordonnées de σ d'indices appartenant à u_2.

Il existe a positif , m entier strictement plus grand que 1 et q entier pair vérifiant : $m.q > p$ et

$$\sup_{\sigma \in \Sigma} \frac{1}{1+\|\sigma\|^a} \int |\Delta_u g^{\frac{1}{q}}(x,\sigma)|^q \, d\lambda_r(x) < \infty$$

pour toutes les combinaisons u satisfaisant : $\ell(u_1) + \ell(u_2) \leq m$.

C : pour tout R positif, il existe une constante b(R) strictement positive telle que :

$$\forall \sigma_o, \|\sigma_o\| \leq R , \overline{\lim_{\|(\mu,\sigma)\| \to \infty}} \|(\mu,\sigma)\|^b \ \rho\left[(0,\sigma_o) ; (\mu,\sigma_o+\sigma)\right] < \infty$$

où ρ désigne l'affinité de Hellinger :

$$\rho\left[(0,\sigma_o) ;(\mu,\sigma_o+\sigma)\right] = \int g^{1/2}(x,\sigma_o)\, g^{1/2}(x-\mu,\ \sigma_o+\sigma) d\lambda_r(x)$$

D : pour tout ε positif ,

$$\lim_{M\to\infty} P_{\sigma_o} \left\{ \sup_{\|(\mu,\sigma)\| \geq M} \frac{g(X-\mu ;\ \sigma_o+\sigma)}{g(X ; \sigma_o)} > \varepsilon \right\} = 0$$

E : La matrice Hessienne $\ddot{L}(x,\sigma)$ de Log $g(x,\sigma)$ en x et σ est continue par rapport aux deux arguments et il existe δ_o positif tel que :

$$E_{\sigma_o} \sup_{\|(\mu,\sigma)\| < \delta_o} \|\ddot{L}(X-\mu,\ \sigma_o+\sigma)\| < \infty$$

F : Conditions sur f :

Il existe des constantes γ et Γ strictement positives telles que :

quels que soit t_1 et t_2 dans $[0,1]$, $t_1 < t_2$, on peut trouver t_3 et t_4 tels que :

$$[t_3,t_4] \subset [t_1,t_2] , \quad t_4-t_3 > \gamma(t_2-t_1)$$

et $\forall\, \eta \in \mathbb{R}^{p_1}$:

$$\inf_{s\in[t_3,t_4]} \|f(s)\eta\| \geqslant \frac{1}{\Gamma} \sup_{s\in[t_1,t_2]} \|f(s)\eta\|$$

Remarques :

i) Les hypothèses A à E sont celles posées dans [21] pour montrer le théorème d'invariance correspondant à l'exemple 1. Seule la condition B se trouve modifiée par l'adjonction de combinaison avec répétitions nécessaires pour la partie régression du modèle.

ii) La condition F servira en particulier à montrer que les variables $\xi_{n,i} = \mathrm{Log} \dfrac{g(X_i - f(\frac{i}{n})\frac{\eta}{\sqrt{n}}, \sigma_o + \frac{\sigma}{\sqrt{n}})}{g(x_i, \sigma_o)}$ sont uniformément asymptotiquement négligeables (u.a.n). Elle est remplie dès que f est lipschitzienne sur $[0,1]$.

Notons $\mathcal{F}(s)$ la matrice $\left[\begin{array}{c|c} f(s) & 0 \\ \hline 0 & I_{p_2 x p_2} \end{array}\right]$, $\overline{F}(t)$ l'intégrale $\int_0^t \mathcal{F}(s)'\,\mathcal{F}(s)ds$ et $\mathcal{J}_{[o,t]}(\sigma_o)$ l'intégrale $\int_0^t \mathcal{F}(s)'\,\mathcal{J}(\sigma_o)\,\mathcal{F}(s)ds$.

Théorème 1 : Sous les conditions A à F, sous l'hypothèse $H_{on}(\theta_o)$ la suite de processus $Z_n^f(\cdot,\cdot)$ converge étroitement dans $\mathcal{C}_\phi(\mathbb{R}^p \times [0,1])$ uniformément en σ_o, pour $\|\sigma_o\| < R$, vers le processus $Z^f(\cdot,\cdot)$

$$Z^f(\theta,t) = \exp \frac{1}{q}\left\{ \theta' \int_0^t \mathcal{F}(s)'\mathcal{J}^{\frac{1}{2}}(\sigma_o)dW(s) - \frac{1}{2}\,\theta'\mathcal{J}_{[o,t]}(\sigma_o)\theta\right\}$$

où W est une Brownien réel à $r + p_2$ dimensions, dès que la fonction matricielle Φ vérifie la condition (1.0) :

(1.0) $\quad \exists\, \alpha > 0 \quad \exists\, c > 0, \quad \forall\, t \in [0,1] \quad : \Phi(t) \leq c\, t^{\alpha} \bar{F}(t)$

Nous démontrons la convergence du processus de vraisemblance lui-même car elle a lieu dans un espace plus adapté au contrôle à l'infini ($\mathcal{C}_o$ ou $\mathcal{C}_\Phi$) ; mais il est immédiat d'en déduire la convergence uniforme sur les compacts de son logarithme car le processus limite ne charge pas 0.

La démonstration de ce théorème fait l'objet du paragraphe IV, ainsi que la proposition suivante qui se démontre avec les mêmes techniques.

Proposition (1.1)

Sous les hypothèses A à F et sous l'hypothèse supplémentaire

$\underline{\underline{H}}$: il existe $k_o \in \mathbb{N}$, $\alpha > 0$ et $\Delta_o > 0$, tels que :

$$E_{\sigma_o} \sup_{(\mu,\sigma)\in\mathbb{R}^{r+p_2}} \prod_{i=1}^{k} \sup_{(\mu',\sigma'),\, \|(\mu',\sigma')-(\mu,\sigma)\| \leq \Delta_o} \left| \frac{g(X_i-\mu',\sigma_o+\sigma')}{g(X_i,\sigma_o)} \right|^{\alpha} < \infty \,.$$

Pour toute fonction $\varphi(\cdot)$ continue sur $[0,1]$ satisfaisant : $\varphi(t) \leqslant C.t^{\beta}$ avec $\beta > 0$,

la suite de processus de $\mathscr{C}[0,1]$ définis par

$$\tau_n(t) = \varphi(t) \sup_{\theta\in\mathbb{R}^p} \operatorname{Log} Z_n^f(\theta,t)$$

converge étroitement sous $H_{o,n}(\theta_o)$ et uniformément en σ_o pour $|\sigma_o| \leq R$, vers le processus :

$$\tau(t) = \varphi(t) \sup_{\theta\in\mathbb{R}^p} \operatorname{Log} Z^f(\theta,t)$$

Corollaire (1.1)

Sous les mêmes hypothèses, si A_n est une suite de réels satisfaisant $A_n > c.n^{\beta}$ avec $\beta > 0$. Alors sous $H_{o,n}(\theta_o)$ et uniformément en σ_o pour $|\sigma_o| < R$. La suite de processus

$\frac{1}{A_n} \log Z_n^f(\cdot,\cdot)$ converge vers 0 uniformément sur $[0,1] \times \mathbb{R}^p$.

La démonstration du corollaire est immédiate en prenant dans la proposition (1.1) $\varphi(\frac{k}{n}) = \frac{1}{c}(\frac{k}{n})^{\frac{\beta}{2}}$.

Pour énoncer la proposition suivante qui est un corollaire trivial du théorème 1 nous remarquons que l'hypothèse $H_{1,n}(t_o,\theta^*,\theta^{**})$ peut être paramétrisée, soit par le triplet $(t_o,\theta^*,\theta^{**})$ soit par le triplet $(t_o, \bar{\theta}, d)$ où $\bar{\theta} = \mathcal{I}_{[o,t_o]}(\sigma_o)\theta^* + (\mathcal{I}_{[o,1]}(\sigma_o) - \mathcal{I}_{[o,t_o]}(\sigma_o))\theta^{**}$ et $d = \theta^{**} - \theta^*$. Ce paramétrage donne des expressions plus commodes et s'interprète facilement : d amplitude du changement , $\bar{\theta}$ paramètre de "centralité". Soit $\bar{\Lambda}_n$ la suite de processus de $\mathcal{C}([0,1]\times\mathbb{R}^{2p})$ définie par interpolation linéaire sur les points

$$(\frac{k}{n},\bar{\theta},d) \longrightarrow \bar{\Lambda}_n(\frac{k}{n},\bar{\theta},d) = \left[\frac{dP_{H_{1n}}(\frac{k}{n},\ \bar{\theta}_o + \frac{\bar{\theta}}{\sqrt{n}}\ ,\ \frac{d}{\sqrt{n}})}{dP_{H_{on}}(\bar{\theta}_o)}\right]^{1/q}$$

Proposition (1.2)

Supposons vérifiées les conditions A à F. Sous l'hypothèse $H_{on}(\bar{\theta}_o)$ la suite de processus $\bar{\Lambda}_n$ converge étroitement dans $\mathcal{C}_\Phi([0,1]\times\mathbb{R}^{2p})$ uniformément en $\sigma_o,(\|\sigma_o\| \leq R)$, vers le processus $\bar{\Lambda}$ défini par la relation

$$\bar{\Lambda}(t,\bar{\theta},d) = \exp\frac{1}{q}\{\bar{\theta}'\mathcal{I}_{[01]}^{-1}(\sigma_o)X(1) + d'(\mathcal{I}_{[0t]}(\sigma_o)\mathcal{I}_{[01]}^{-1}(\sigma_o)X(1) - X(t))$$

$$- \frac{1}{2}\ \bar{\theta}'\mathcal{I}_{[01]}(\sigma_o)\bar{\theta}$$

$$- \frac{1}{2}\ d'\mathcal{I}_{[0t]}(\sigma_o)(I_{pp} - \mathcal{I}_{[01]}^{-1}(\sigma_o)\ \mathcal{I}_{[0t]}(\sigma_o))d\}$$

avec $X(t) = \int_o^t \mathcal{F}(s)'\ \mathcal{I}^{1/2}(\sigma_o)dW(s)$,

dès que la fonction matricielle Φ vérifie la condition (1.1)

$$\exists\alpha > 0,\ \exists c > 0\ /\ \forall t \in [0,1] \qquad (1.1)$$
$$\phi(1) - c(1-t)^\alpha\ (\bar{F}(1) - \bar{F}(t)) \leq \Phi(t) \leq ct^\alpha\ \bar{F}(t)$$

II.- TESTS DE L'EXISTENCE D'UNE RUPTURE

Dans ce paragraphe, nous allons étudier le comportement du test de vraisemblance. En effet cette procédure semble avoir eu les préférences de diverses comparaisons effectuées parmi les procédures classiques. Plus précisemment, pour ce qui est de la comparaison asymptotique locale, l'expérience limite n'étant pas à rapport de vraisemblance monotone par rapport au paramètre t d'instant de rupture (t appartenant à $]0,1[$), aucun test n'est sélectionnable (voir [19]) ; par la comparaison asymptotique non locale (voir [18] et [20]) on montre que dans un large éventail de modèles, le test du rapport de vraisemblance est optimal ; enfin sur un modèle gaussien unidimensionnel, la comparaison par des méthodes de Monte-Carlo (voir [20]) semblent confirmer largement ce dernier résultat.

Toutefois, nous amènerons ici quelques restrictions : tout d'abord dans la proposition (2.1), nous montrons qu'à condition d'exclure strictement les bords ($t = 0$ et $t = 1$), on peut approximer le niveau du test de vraisemblance mais dans la proposition (2.2), on montre que cela n'est plus possible si on n'exclut pas les bords.

La dernière partie de ce paragraphe est consacrée à une remormalisation de la statistique de vraisemblance de manière à faire intervenir les bords de façon pondérée.

En revenant au premier paramétrage (θ^*,θ^{**}) qui est plus adapté ici, notons $\Lambda_n(k, \theta^*, \theta^{**} ; \theta)$ le rapport de vraisemblance (à valeurs dans $[1,\infty[$) de l'hypothèse alternative $H_{1,n}(k,\theta_o+\frac{\theta^*}{\sqrt{n}},\theta_o+\frac{\theta^{**}}{\sqrt{n}})$ à l'hypothèse nulle $H_{o,n}(\frac{\theta}{\sqrt{n}}+\theta_o)$ et $\Lambda(t, \theta^*, \theta^{**} ; \theta)$ celui correspondant au processus limite (s'exprimant sans confusion possible à partir de $\overline{\Lambda}(t ; \overline{\theta},d)$)

$$\Lambda(t, \theta^*, \theta^{**} ; \theta) = \left[\frac{Z(\theta^*;t)\, Z(\theta^{**};1)}{Z(\theta^{**};t) Z(\theta ; 1)}\right]^q$$

Dans les paragraphes II et III, nous supprimerons les indices f dans les processus Z_n^f et Z^f, la confusion n'étant pas possible.

Proposition (2.1)

Sous les conditions A à F , pour tout $\varepsilon \in]0, \frac{1}{2}]$, pour tout A supérieur ou égal à 1 ,

$$\lim_{n\to\infty} P_{H_{o,n}(\theta_o)} \{ \sup_{n\varepsilon\leqslant k\leqslant n(1-\varepsilon)} \sup_{\theta^*,\theta^{**}\in\mathbb{R}^p} \inf_{\theta\in\mathbb{R}^p} \Lambda_n(k, \theta^*, \theta^{**} ; \theta) > A\}$$

$$= P_{\sigma_o} \{ \sup_{\varepsilon\leqslant t\leqslant 1-\varepsilon} \sup_{\theta^*,\theta^{**}\in\mathbb{R}^p} \inf_{\theta\in\mathbb{R}^p} \Lambda(t, \theta^*, \theta^{**} ; \theta) > A\} \qquad (2.0)$$

$$= P\{ \sup_{\varepsilon\leqslant t\leqslant 1-\varepsilon} \left[\int_o^t \mathscr{F}(s)' \mathcal{J}^{1/2}(\sigma_o) dW(s)\right]' \mathcal{J}^{-1}_{[o,t]}(\sigma_o) \left[\int_o^t \mathscr{F}(s)' \mathcal{J}^{1/2}(\sigma_o) dW(s)\right]$$

$$+ \left[\int_t^1 \mathscr{F}(s)' \mathcal{J}^{1/2}(\sigma_o) dW(s)\right]' \mathcal{J}^{-1}_{[t,1]}(\sigma_o) \left[\int_t^1 \mathscr{F}(s)' \mathcal{J}^{1/2}(\sigma_o) dW(s)\right]$$

$$- \left[\int_o^1 \mathscr{F}(s)' \mathcal{J}^{1/2}(\sigma_o) dW(s)\right]' \mathcal{J}^{-1}_{[0,1]}(\sigma_o) \left[\int_o^1 \mathscr{F}(s)' \mathcal{J}^{1/2}(\sigma_o) dW(s)\right]$$

$$\geqslant \text{Log } A\}$$

où W désigne un mouvement brownien à $r+p_2$ dimensions sur $[0,1]$.

Démonstration

Remarquons d'abord que la seconde égalité résulte d'un calcul facile sur le processus limite Λ et montrons donc l'égalité (2.0). Le théorème (1.1) montre la convergence dans $\mathscr{C}_\Phi$ de la suite de processus $Z_n(\theta,t)$; on va en tirer aisément la convergence de la suite

$$\Lambda_n(k ; \theta^*, \theta^{**} ; \theta) = \left[\frac{Z_n(\theta^* ; \frac{k}{n}) \cdot Z_n(\theta^{**} ; 1)}{Z_n(\theta^{**} ; \frac{k}{n}) \cdot Z_n(\theta ; 1)}\right]^q$$

à l'aide du lemme suivant (trivial mais utilisé souvent dans la suite).

Lemme (2.1)

Sous les conditions du théorème (1.1), pour tous ε et $\varepsilon' > 0, (\varepsilon < 1)$, il existe M satisfaisant pour tout n :

$$P\{ \sup_{t\in[\varepsilon,1]} \sup_{\substack{\theta \\ t\theta\phi(t)\theta\leq M}} Z_n(\theta,t) = \sup_{t\in[\varepsilon,1]} \sup_{\theta\in\mathbb{R}^p} Z_n(\theta,t)\} \geq 1-\varepsilon' \qquad (2.1)$$

Démonstration du lemme

Il suffit de remarquer que pour ε' fixé ($\varepsilon' < 1$), il existe M satisfaisant :

$$P\{ \sup_{\substack{t \\ t\theta\phi(t)\theta>M}} \sup_{\theta} Z_n(\theta,t) > \varepsilon'\} < \varepsilon' \qquad (2.2)$$

et que $Z_n(0,t) \equiv 1$.

Nous appliquons ce lemme avec $\phi(t) = I_{p\times p}.1_{[\varepsilon,1-\varepsilon]}(t)$, ce qui implique l'existence de $M(\varepsilon,\varepsilon')$ vérifiant pour tout n :

$$P\{ \sup_{t\in[\varepsilon,1-\varepsilon]} \sup_{\theta^*,\|\theta^*\|\leq M} Z_n(\theta^*,t) \neq \sup_{t\in[\varepsilon,1-\varepsilon]} \sup_{\theta^*\in\mathbb{R}^p} Z_n(\theta^*,t)\} \leq \varepsilon'$$

et aussi

$$P\{ \sup_{t\in[\varepsilon,1-\varepsilon]} \sup_{\theta^{**},\|\theta^{**}\|\leq M} \frac{Z_n(\theta^{**},1)}{Z_n(\theta^{**},t)} \neq \sup_{t\in[\varepsilon,1-\varepsilon]} \sup_{\theta^{**}\in\mathbb{R}^p} \frac{Z_n(\theta^{**};1)}{Z_n(\theta^{**};t)}\} \leq \varepsilon'$$

car les processus $\dfrac{Z_n(\theta^{**};1)}{Z_n(\theta^{**};t)}$ et $\overleftarrow{Z}^f_n(\theta^{**};1-t)$ ont même loi, avec $\overleftarrow{f}(t) = f(1-t)$. Une conséquence encore plus triviale en est :

$$P\{ \sup_{\theta,\|\theta\|\leq M} Z_n(\theta\,;1) \neq \sup_{\theta\in\mathbb{R}^p} Z_n(\theta\,;1)\} \leq \varepsilon'$$

Ces trois dernières inégalités permettent d'approcher à $3\varepsilon'$ près, pour tout n, la probabilité :

$$P_{H_{o,n}(\theta_o)}\{\sup_{n\varepsilon\leq k\leq n(1-\varepsilon)}\ \sup_{\theta^*,\theta^{**}\in\mathbb{R}^p}\ \inf_{\theta\in\mathbb{R}^p}\ \Lambda_n(k,\theta^*,\theta^{**};\theta) > A\}$$

par $P_{H_{o,n}(\theta_o)}\{\sup_{n\varepsilon\leq k\leq n(1-\varepsilon)}\ \sup_{\theta^*,\|\theta^*\|\leq M}\ \sup_{\theta^{**},\|\theta^{**}\|\leq M}\ \inf_{\theta,\|\theta\|\leq M}\ \Lambda_n(k,\theta^*,\theta^{**},\theta) > A\}$

et cette dernière quantité converge, à M fixé, vers

$$P_{\sigma_o}\{\sup_{\varepsilon\leq t\leq 1-\varepsilon}\ \sup_{\theta^*,\|\theta^*\|\leq M}\ \sup_{\theta^{**},\|\theta^{**}\|\leq M}\ \inf_{\theta,\|\theta\|\leq M}\ \Lambda(t,\theta^*,\theta^{**};\theta) > A\}$$

enfin cette probabilité diffère de moins de $3\varepsilon'$ (même argument que ci-dessus) de la limite cherchée :

$$P_{\sigma_o}\{\sup_{\varepsilon\leq t\leq 1-\varepsilon}\ \sup_{\theta^*,\theta^{**}\in\mathbb{R}^p}\ \inf_{\theta\in\mathbb{R}^p}\ \Lambda(t,\theta^*,\theta^{**};\theta) > A\}$$

ce qui achève la preuve de la proposition (2.1).

L'expression du processus limite en fonction du mouvement brownien donne :

$$\sup_{t\in[\varepsilon,1-\varepsilon]}\ \sup_{\theta^*,\theta^{**}\in\mathbb{R}^p}\ \inf_{\theta\in\mathbb{R}^p}\ \Lambda(t,\theta^*,\theta^{**};\theta)$$

$$= \sup_{t\in[\varepsilon,1-\varepsilon]} X(t)'\,\mathcal{J}^{-1}_{[0,t]}\,X(t) + \left[X(1)-X(t)\right]'\mathcal{J}^{-1}_{[t,1]}\left[X(1)-X(t)\right] - X(1)'\mathcal{J}^{-1}_{[0,1]}X(1)$$

avec $X(t) = \int_0^t \mathcal{F}(s)'\,\mathcal{J}^{1/2}(\sigma_o)\,dW(s)$

La loi du logarithme itéré pour le mouvement brownien implique la convergence de : $\sup_{t\in[\varepsilon,1]} X(t)'\,\mathcal{J}^{-1}_{[0,t]}\,X(t)$ vers l'infini lorsque ε tend vers 0 et de même celle de $\sup_{t\in[0,1-\varepsilon]} \left[X(1)-X(t)\right]'\mathcal{J}^{-1}_{[t,1]}\left[X(1)-X(t)\right]$ et par suite la :

Proposition (2.2)

Sous les conditions A à F , pour tout A positif, on a :

$$\lim_{n\to\infty} P_{H_{o,n}(\theta_o)}\{ \sup_{1\leqslant k\leqslant n} \sup_{\theta^*,\theta^{**}\in\mathbb{R}^p} \inf_{\theta\in\mathbb{R}^p} \Lambda_n(k, \theta^*, \theta^{**}; \theta) > A\} = 1 .$$

Sans modifications, le test du rapport de vraisemblance (avec $t\in\,]0,1[$) n'est donc pas de niveau approximable. On pourrait penser à multiplier Λ_n par une suite de constantes C_n convergeant vers 0 de sorte que la statistique de test ne converge plus vers l'infini mais la proposition (2.1) met en évidence la perte de toute sensibilité du test pour une rupture n'intervenant pas aux bords. La seule possibilité qui nous reste est de multiplier Λ_n (ou son logarithme) par une constante $C_n(k)$ variant avec k, c'est-à-dire de modifier la vraisemblance en la pénalisant aux bords.

Comme on l'a vu dans [19] pour le cas particulier sans régression : $p_1 = 0$, $p_2 = 1$, une solution possible consiste à multiplier le Log-vraisemblance : $\text{Log}\,\Lambda_n(k,\cdot,\cdot;\cdot)$ par $2\,\frac{k}{n}(1-\frac{k}{n})$.

Le corollaire de la proposition (1.1) suivant permet de préciser les possibilités d'une renormalisation de la log-vraisemblance :

Corollaire (2.1)

Sous les hypothèses de la proposition 1.1. . Pour toute fonction réelle φ continue sur $[0,1]$ et vérifiant la condition (2.1) :

$$\exists\beta>0, \exists C>0 \;/\; \forall t\in[0,1]\;\; \varphi(1) - C(1-t)^{\beta} < \varphi(t) < Ct^{\beta}$$

la suite de variables aléatoires $\{M_n\}$:

$$M_n = \sup_{1\leqslant k\leqslant n} \sup_{\theta^*,\theta^{**}\in\mathbb{R}^p} \inf_{\theta\in\mathbb{R}^p} (\varphi(\tfrac{k}{n})\,\text{Log}\,\Lambda_n(k,\theta^*,\theta^{**},\theta)(\varphi(1)-\varphi(\tfrac{k}{n}))$$

converge sous $H_{on}(\theta_o)$ vers la variable : $\sup_{t\in[0,1]} \varphi(t)B_o(X(t))(\varphi(1)-\varphi(t))$ avec

$$B_o(X(t))=\{ (X(t)-\mathcal{J}_{[o,t]}(\sigma_o)\,\mathcal{J}^{-1}_{[o,1]}(\sigma_o)\,X(1))'(\mathcal{J}^{-1}_{[o,t]}(\sigma_o) + \mathcal{J}^{-1}_{[t,1]}(\sigma_o))$$
$$(X(t)-\mathcal{J}_{[o,t]}(\sigma_o)\,\mathcal{J}^{-1}_{[o,1]}(\sigma_o)X(1))\}$$

Le problème alors consiste à trouver une normalisation φ telle que cette statistique soit libre sous H_o. La réponse à cette question dépend étroitement de la forme de $\mathcal{J}_{[o,t]}(\sigma_o)$. Remarquons cependant que c'est faux en toute généralité, comme le montre le raisonnement suivant:

Supposons qu'il existe φ tel que $\sup_t (\varphi(1)-\varphi(t))\ \varphi(t) B_o(X(t))$ soit indépendant de $\mathcal{J}(\sigma_o)$ pour $\mathcal{J}(\sigma_o)$ appartenant à un sous espace assez grand, il est alors facile de montrer que cela entraîne que la loi

$$\sup_{t\in[0,1]} (\varphi(1)-\varphi(t))\varphi(t) \sum_{i=1}^{r+p_2} \frac{B_i(\lambda_i(t))^2}{\lambda_i(t)(1-\lambda_i(t))}$$

est indépendante des λ_i pour $B_1,\ldots,B_{r+p_2}$, suite de Ponts Browniens réels indépendants et λ_i fonctions croissantes continues de $[0,1]$ dans $[0,1]$ ce qui est faux (en le vérifiant sur des fonctions particulières). Cependant, comme on le verra dans l'étude d'exemples à la fin de ce paragraphe cette normalisation est utilisable dans nombre de cas classiques ($p_1 = 0$, p_2 quelconque ; p_1 quelconque, $p_2 = 0$, exemple 2 bis ...).

Quand ce point de vue n'est pas utilisable, nous proposons alors de faire une renormalisation "interne" à la vraisemblance qui peut pallier à l'inconvénient dû à la régression : elle consiste à :

i) calculer les estimateurs du maximum de vraisemblance $\hat{\theta}^*_{n,k}$ et $\hat{\theta}^{**}_{n,k}$ de θ^* et θ^{**} à k fixé ;

ii) pondérer les estimateurs eux-mêmes par une fonction de k ;

iii) calculer la vraisemblance aux maxima pondérés et prendre le suprémum en k.

En parallèle au corollaire (2.1) pour la renormalisation externe, la proposition suivante (2.3) permettra d'envisager les différentes possibilités de renormalisation interne.

Proposition (2.3)

Soit Ψ une fonction matricielle symétrique à valeurs dans $\mathbb{R}^{p\times p}$, régulière, continue sur $[0,1]$ et telle que $\Phi(t) = \Psi^2(t)$ satisfait la condition (1.0). Alors, si $\hat{\theta}_{n,k}$ désigne pour $k=1,\ldots,n$ un choix mesurable dans l'ensemble des $\hat{\theta}$ réalisant le suprémum de $Z_n(.,\frac{k}{n})$:

i) la famille de vecteurs de $\mathbb{R}^p$: $\{\psi(\frac{k}{n})\hat{\theta}_{n,k},\ k=1,\ldots,n\}_{n\in\mathbb{N}}$ est tendue

ii) la suite de processus $\overset{\vee}{\tau}_n$ de $\mathscr{C}[0,1]$ ayant pour trajectoires les lignes polygonales joignant les points $(0,1)$ et $(\frac{k}{n}, \overset{\vee}{\tau}(\frac{k}{n}))$

$$\overset{\vee}{\tau}_n(\tfrac{k}{n}) = Z_n(\psi(\tfrac{k}{n})\,\hat{\theta}_{n,k}\,;\,\tfrac{k}{n})$$

converge étroitement sous $H_{o,n}(\theta_o)$ vers le processus $\overset{\vee}{\tau}$:

$$\begin{aligned}\overset{\vee}{\tau}(t) &= Z(\psi(t)\hat{\theta}_t\,;\,t)\\ &= \exp\tfrac{1}{q}\Big\{X(t)'\Big[\mathcal{I}^{-1}_{[0,\frac{t}{2}]}(\sigma_o)\psi(t)\\ &\quad - \tfrac{1}{2}\,\mathcal{I}^{-1}_{[0,t]}(\sigma_o)\psi(t)\,\mathcal{I}_{[0,t]}(\sigma_o)\psi(t)\,\mathcal{I}^{-1}_{[0,t]}(\sigma_o)\Big]X(t)\Big\}\end{aligned}$$

Démonstration de la proposition (2.3)

i) la tension de la famille $\{\psi(\frac{k}{n})\hat{\theta}_{n,k}\}$ résulte de celle de la famille $\{\hat{\theta}'_{n,k}\,\phi(\frac{k}{n})\hat{\theta}_{n,k}\}$, la suite de processus Z_n convergeant dans

$$\mathscr{C}_\phi : \forall\varepsilon' > 0\ ,\ \exists M \text{ tel que } P\{\sup_{1\leqslant k\leqslant n}\hat{\theta}'_{n,k}\,\phi(\tfrac{k}{n})\,\hat{\theta}_{n,k} > M\}$$

$$\leqslant P\{\sup_{1\leqslant k\leqslant n}\ \sup_{\substack{\theta\\ \theta'\phi(\frac{k}{n})\theta>M}} Z_n(\theta\,;\tfrac{k}{n}) > 1\} \leqslant \varepsilon' \qquad (2.3)$$

ii) la démonstration est du même type que celle de la proposition (1.1) Nous allons montrer en fait la tension au voisinage de $t = 0$ et la tension sur tout compact inclus dans $]0,1]$ de la suite $\overset{\vee}{\tau}_n$, la convergence des marginales finies étant facile.

Pour tous ε et ε' et M positifs, on peut écrire :

$$P\{ \sup_{o\leqslant k\leqslant n\varepsilon} |Z_n(\psi(\tfrac{k}{n})\hat{\theta}_{n,k}; \tfrac{k}{n}) - 1| > \varepsilon'\}$$

$$\leqslant \{ \sup_{o\leqslant \frac{k}{n}\leqslant\varepsilon} \sup_{\theta,\|\theta\|\leqslant M} |Z_n(\theta,\tfrac{k}{n})-1| > \varepsilon'\} + P\{ \sup_{1\leqslant k\leqslant n} \|\psi(\tfrac{k}{n})\hat{\theta}_{n,k}\| > M\}$$

Fixant ε', le second terme est inférieur à ε' pour $M \geqslant M_o(\varepsilon')$ à cause de (2.3) ; ε' et $M_o(\varepsilon')$ étant fixés, le premier terme est inférieur à ε' pour $\varepsilon \leqslant \varepsilon_o(\varepsilon', M_o(\varepsilon'))$ à cause de la tension en $t = 0$ de la suite $Z_n(\cdot,\theta)$ restreinte à tout compact en θ.

Montrons pour finir la tension dans $\mathscr{C}[\varepsilon_o,1]$ de la suite $\overset{\vee}{\tau}_n$:

$$P\{ \sup_{\substack{n\varepsilon_o\leqslant k\leqslant k'\leqslant n \\ k'\leqslant k+n\delta}} |\overset{\vee}{\tau}_n(\tfrac{k'}{n}) - \overset{\vee}{\tau}_n(\tfrac{k}{n})| > \varepsilon'\}$$

$$\leqslant 2P\{ \sup_{n\varepsilon_o\leqslant k\leqslant n} \|\psi(\tfrac{k}{n})\hat{\theta}_{n,k}\| > M_o\}$$

$$+ P\{ \sup_{\substack{n\varepsilon_o\leqslant k\leqslant k'\leqslant n \\ k'\leqslant k+n\delta}} |Z_n(\psi(\tfrac{k'}{n})\hat{\theta}_{n,k'}; \tfrac{k'}{n}) - Z_n(\psi(\tfrac{k}{n})\hat{\theta}_{n,k};\tfrac{k}{n})| > \varepsilon' \cap \sup_{n\varepsilon_o\leqslant \ell\leqslant n} \|\psi(\tfrac{\ell}{n})\hat{\theta}_{n,\ell}\| \leqslant M_o\}$$

$$< 2\varepsilon' + P\{ \sup_{\substack{n\varepsilon_o\leqslant k\leqslant k'\leqslant n \\ k'\leqslant k+n\delta}} \|\psi(\tfrac{k'}{n})\hat{\theta}_{n,k'} - \psi(\tfrac{k}{n})\hat{\theta}_{n,k}\| > h \cap \sup_{n\varepsilon_o\leqslant \ell\leqslant n} \|\psi(\tfrac{\ell}{n})\hat{\theta}_{n,\ell}\| \leqslant M_o\}$$

$$+ P\{ \sup_{\substack{n\varepsilon_o\leqslant k\leqslant k'\leqslant n \\ k'\leqslant k+n\delta}} \sup_{\theta,\|\theta\|\leqslant M_o} \sup_{\substack{\theta^1,\ \|\theta^1\|\leqslant M_o \\ \|\theta-\theta^1\|\leqslant h}} |Z_n(\theta^1, \tfrac{k'}{n}) - Z_n(\theta, \tfrac{k}{n})| > \varepsilon'\}$$

Ce dernier terme est inférieur à ε' pour tout n dès que $\delta < \delta_1(\varepsilon')$ et

$h < h_o(\varepsilon')$ à cause de la tension de la suite $Z_n(\cdot,\cdot)$. Il reste donc à majorer :

$$P\{ \sup_{\substack{n\varepsilon_o \leq k \leq k' \leq n \\ k' \leq k+n\delta}} \| \psi(\tfrac{k'}{n}) \hat{\theta}_{n,k'} - \psi(\tfrac{k}{n}) \hat{\theta}_{n,k} \| > h_o \cap \sup_{n\varepsilon_o \leq \ell \leq n} \| \psi(\tfrac{\ell}{n}) \hat{\theta}_{n,\ell} \| \leq M_o \}$$

La continuité de ψ permet de majorer par ε' dès que $\delta < \delta_2(\varepsilon')$ la quantité

$$P\{ \sup_{\substack{n\varepsilon_o \leq k \leq k' \leq n \\ k' \leq k+n\delta}} \| [\psi(\tfrac{k'}{n}) - \psi(\tfrac{k}{n})] \hat{\theta}_{n,k'} \| > \frac{h_o}{2} \cap \sup_{n\varepsilon_o \leq \ell \leq n} \| \psi(\tfrac{\ell}{n}) \hat{\theta}_{n,\ell} \| \leq M_o \}$$

Il reste par conséquent à montrer que pour δ petit on a :

$$P\{ \sup_{\substack{n\varepsilon_o \leq k \leq k' \leq n \\ k' \leq k+\delta n}} \| \hat{\theta}_{n,k} - \hat{\theta}_{n,k'} \| > \frac{h_o}{2 \| \psi(1) \|} \cap \sup_{n\varepsilon_o \leq \ell \leq n} \| \hat{\theta}_{n,\ell} \| \leq M_o \| \psi^{-1}(\varepsilon_o) \| \} < \varepsilon'$$

Notons $h'_o = \dfrac{h_o}{2 \| \psi(1) \|}$ et $M'_o = M_o \| \psi^{-1}(\varepsilon_o) \|$ et utilisons de nouveau les propriétés de continuité et de tension de la suite Z_n : pour tout A positif, on a :

$$P\{ \sup_{\substack{n\varepsilon_o \leq k \leq k' \leq n \\ k' \leq k+n\delta}} \| \hat{\theta}_{n,k} - \hat{\theta}_{n,k'} \| > h'_o \cap \sup_{n\varepsilon_o \leq \ell \leq n} \| \hat{\theta}_{n,\ell} \| \leq M'_o \}$$

$$< P\{ \sup_{\substack{\varepsilon_o \leq t \leq t' \leq 1 \\ t' \leq t+\delta}} \Big| \sup_{\theta, \|\theta\| \leq M'_o} Z_n(\theta;t) - \sup_{\theta', \|\theta'\| \leq M'_o} Z_n(\theta';t') \Big| \geq A \}$$

$$+ P\{ \sup_{\substack{\varepsilon_o \leq t \leq t' \leq 1 \\ t' \leq t+\delta}} \sup_{\theta, \|\theta\| \leq M'_o} |Z_n(\theta,t) - Z_n(\theta,t')| \geq A \}$$

$$+ P\{ \sup_{\varepsilon_o \leq t \leq 1} \sup_{\theta, \|\theta\| \leq M'_o} \Big[\sup_{\substack{\tilde{\theta}, \|\tilde{\theta}\| < M'_o \\ \|\tilde{\theta}-\theta\| < h'_o}} Z_n(\tilde{\theta};t) - \sup_{\substack{\nu\theta, \|\nu\theta\| \leq M'_o \\ \|\nu\theta-\theta\| \geq h'_o}} Z_n(\nu\theta,t) \Big] \leq 2A \}$$

Le dernier terme est inférieur à $\dfrac{\varepsilon'}{3}$ pour $A \leq A_o(\varepsilon', h'_o, M'_o)$ parce que

la suite Z_n^* converge vers Z^* dans $\mathscr{C}([0,1]\times\mathbb{R}^p)$ et que la majoration est vraie pour le processus Z. Ayant maintenant fixé $A = A_o$, les deux premiers termes sont aussi inférieurs à $\frac{\varepsilon'}{3}$ dès que δ est inférieur à $\delta_3(\varepsilon')$.

Comme pour la renormalisation externe, se pose le problème de l'existence de Φ normalisante qui aboutisse à une statistique de test libre sous H_o : on peut montrer de même qu'en toute généralité c'est faux et plus précisément, nous proposons une solution au problème dans deux cas : celui où l'information $\mathcal{I}(\sigma)$ est connue à une constante multiplicative près : $\mathcal{I}(\sigma) = i(\sigma).\mathcal{I}$ et dans celui où les paramètres η et σ se séparent asymptotiquement : la matrice $\mathcal{I}(\sigma)$ est diagonale par bloc. Nous présentons ici 2 types de renormalisations internes : La première paraît plus naturelle cependant que la seconde conduit à une loi limite plus classique et présente l'avantage d'opérer sur la vraisemblance une modification analogue à celle correspondant à la renormalisation externe (que les utilisateurs s'accordent à trouver naturelle) dans le cas unidimentionnel (cas où les 2 situations peuvent s'envisager simultanément).

En tout état de cause, comme il n'existe pas de test optimal, aussi bien à distance finie qu'asymptotiquement (au sens local) ([20]), c'est le contexte qui indiquera une normalisation pour accroître ou réduire la sensibilité aux bords. On peut remarquer aussi que les théorèmes précédents permettraient tout aussi bien d'employer une procédure teintée de bayésien qui maximiserait l'espérance par rapport à une loi à priori sur t du rapport de vraisemblance au point t (dont on a préalablement pris les suprema-infima en θ^*, θ^{**}, θ).

Nous considérons les deux normalisations définies à l'aide de

$$\psi(\tfrac{k}{n}) = \tilde{\mathcal{I}}_{[0,\frac{k}{n}]} \qquad \text{et} \qquad \psi^1(\tfrac{k}{n}) = 2\,\tilde{\mathcal{I}}_{[0,\frac{k}{n}]}$$

avec $\tilde{\mathcal{I}}_{[0,t]} = \int_0^t \mathcal{F}(s)'\mathcal{I}\,\mathcal{F}(s)\,ds$

qui fournissent les statistiques de tests

$$\mathcal{T}_n(\tfrac{k}{n}) = \Lambda_n(k\,;\ \widetilde{\mathfrak{I}}_{[o,\frac{k}{n}]}\ \hat{\theta}^{*}_{n,k}\,;\ \widetilde{\mathfrak{I}}_{[\frac{k}{n},1]}\ \hat{\theta}^{**}_{n,k}\,;\ \widetilde{\mathfrak{I}}_{[0,1]}\ \hat{\theta}_{n,n})$$

$$\mathcal{T}^{1}_n(\tfrac{k}{n}) = \Lambda_n(k\,;\ 2\ \widetilde{\mathfrak{I}}_{[o,\frac{k}{n}]}\ \hat{\theta}^{*}_{n,k}\,;\ 2\ \widetilde{\mathfrak{I}}_{[\frac{k}{n},1]}\ \hat{\theta}^{**}_{n,k}\,;\ 2\ \widetilde{\mathfrak{I}}_{[o,\frac{k}{n}]}\ \hat{\theta}_{n,n})$$

dont les niveaux peuvent être approximés dans le

cas où l'information $\mathfrak{I}(\sigma)$ est connue à une constante multiplicative près ($\mathfrak{I}(\sigma) = i(\sigma)\mathfrak{J}$) : un simple changement de paramétrage permet de se ramener à

$$\int_0^1 \mathcal{F}(s)'\mathfrak{J}\mathcal{F}(s)ds = I_{p,p} \quad \text{d'où } \mathfrak{J}_{[0,1]}(\sigma) = i(\sigma)\, I_{pp}\ ,$$

l'homogénéité en $\mathfrak{J}$ faisant disparaître le $i(\sigma)$ inconnu.

Corollaire (2.2)

Sous les hypothèses de la proposition 1.1 et l'hypothèse supplémentaire l'information est "connue" avec $\mathfrak{I}(\sigma) = i(\sigma)\, I_{p,p}$, on a : $\forall\, A > 0$

$$\text{i)} \quad \lim_{n\to\infty} P_{H_{o,n}(\theta_o)}\{\sup_{1\leq k\leq n} \mathcal{T}_n(\tfrac{k}{n}) > A\}$$

$$= P\{\sup_{t\in[0,1]} \tilde{X}(t)'\Big[I - \frac{1}{2}\,\widetilde{\mathfrak{J}}_{[o,t]}\Big]\tilde{X}(t) + \big[\tilde{X}(1)-\tilde{X}(t)\big]'\Big[I-\frac{1}{2}\,\widetilde{\mathfrak{J}}_{[t,1]}\Big]\big[\tilde{X}(1)-\tilde{X}(t)\big]$$

$$-\frac{1}{2}\,\tilde{X}(1)'\tilde{X}(1) > \text{Log } A\}$$

$$\text{ii)} \quad \lim_{n\to\infty} P_{H_{o,n}(\theta_o)}\{\sup_{1\leq k\leq n} \mathcal{T}^{1}_n(\tfrac{k}{n}) > A\}$$

$$= P\{\sup_{t\in[0,1]} \big[\tilde{X}(t) - \widetilde{\mathfrak{J}}_{[o,t]}\tilde{X}(1)\big]'\big[\tilde{X}(t) - \widetilde{\mathfrak{J}}_{[o,t]}\tilde{X}(1)\big] > \text{Log } A\}$$

où $X(t) = \int_0^t f(s)\, dW(s)$.

Remarque : La normalisation avec $\mathcal{T}^{2}_n$ définie de manière analogue à $\mathcal{T}^{1}_n$ par

$$\mathcal{T}^{2}_n(\tfrac{k}{n}) = \Lambda_n(k\,;\ 2\ \widetilde{\mathfrak{I}}_{[o,\frac{k}{n}]}\ \hat{\theta}^{*}_{n,k}\,;\ 2\ \widetilde{\mathfrak{I}}_{[\frac{k}{n},1]}\ \hat{\theta}^{**}_{n,k}\,;\ 2\ \widetilde{\mathfrak{I}}_{[\frac{k}{n},1]}\ \hat{\theta}_{n,n})$$

conduit au même niveau asymptotique que $\mathcal{T}^{1}_n$.

Précisons les résultats obtenus sur les différents exemples :

Exemple 1 : Test de rupture dans un modèle statistique classique : $g(\cdot,\theta)$.

Proposition (2.4) Sous les hypothèses de la proposition 1.1.

Le test de région de rejet :

$$\sup_{1\leqslant k\leqslant n} \sup_{\theta^{*}\in\mathbb{R}^p} \sup_{\theta^{**}\in\mathbb{R}^p} \inf_{\theta\in\mathbb{R}^p} 2.\frac{k}{n}\left(1-\frac{k}{n}\right) \operatorname{Log} \Lambda_n(k,\theta^{*},\theta^{**};\theta) \geqslant A$$

est de niveau approximable par :

$$P\{ \sup_{t\in[0,1]} \sum_{i=1}^{p} B^{(i)}(t)^2 \geqslant A\}$$

où les $B^{(i)}$ sont p ponts browniens indépendants.

Exemple 2 : Régression-translation sans paramètre auxiliaire : $g(\cdot-f(\cdot)\eta)$. Quitte à faire un changement de paramétrage sur f on supposera $\mathcal{I}_{[0,1]} = I$.

Proposition (2.5)

Sous les hypothèses de la proposition 1.1, les tests de régions de rejet

$$\{ \sup_{1\leqslant k\leqslant n} \Lambda_n\left(k,\ 2\ \mathcal{I}_{\left[0,\frac{k}{n}\right]}\hat{\theta}^{*}_{n,k},\ 2\left[I-\mathcal{I}_{\left[0,\frac{k}{n}\right]}\right]\hat{\theta}^{**}_{n,k};\ 2\ \mathcal{I}_{\left[0,\frac{k}{n}\right]}\hat{\theta}_{n,n}\right) > A\}$$

et $\{ \sup_{1\leqslant k\leqslant n} \Lambda_n\left(k,\ 2\ \mathcal{I}_{\left[0,\frac{k}{n}\right]}\hat{\theta}^{*}_{n,k},\ 2\left[I-\mathcal{I}_{\left[0,\frac{k}{n}\right]}\right]\hat{\theta}^{**}_{n,k},\ 2\left[I-\mathcal{I}_{\left[0,\frac{k}{n}\right]}\right]\hat{\theta}_{n,n}\right) > A\}$

ont un niveau approximable par :

$$P\{ \sup_{0\leqslant t\leqslant 1} \left[X(t) - \mathcal{I}_{[0,t]}\, X(1)\right]' \left[X(t) - \mathcal{I}_{[0,t]}\, X(1)\right] > \operatorname{Log} A\} .$$

Exemple 2 bis : modèle de translation-homothétie : $\frac{1}{\sigma} g(\frac{\cdot - f(.)\eta}{\sigma})$.

L'information $\mathcal{J}(\sigma)$ est connue à la constante multiplicative $\frac{1}{\sigma^2}$ près : $\mathcal{J}(\sigma) = \frac{1}{\sigma^2}\tilde{\mathcal{J}}$

Proposition (2.5 bis)

Sous les hypothèses de la prop. 1.1, et si la fonction de régression f a été orthonormalisée par : $\tilde{\mathcal{J}}_{[o,1]} = I_{p_1,p_1}$, les tests de régions de rejet :

$$\{ \sup_{1\le k\le n} \mathcal{J}^1_n (\tfrac{k}{n}) > A\}$$

et

$$\{ \sup_{1\le k\le n} \mathcal{J}^2_n (\tfrac{k}{n}) > A\}$$

ont un niveau approximable par :

$$P\{ \sup_{0\le t\le 1} [\int_0^t \mathcal{F}(s)' \mathcal{J}^{\frac{1}{2}} dW(s) - \tilde{\mathcal{J}}_{[0,t]} \int_0^1 \mathcal{F}(s)' \mathcal{J}^{\frac{1}{2}} dW(s)]' [\int_0^t \mathcal{F}(s) \mathcal{J}^{\frac{1}{2}} dW(s) - \tilde{\mathcal{J}}_{[0,t]} \int_0^1 \mathcal{F}(s) \mathcal{J}^{\frac{1}{2}} dW(s)] > \log A \}$$

Exemple 3 : Régression translation avec paramètre inconnu supplémentaire : $g(\cdot - f(\cdot)\eta;\sigma)$ dans le cas où la matrice d'information de Fisher est diagonale par bloc : $\mathcal{J}(\sigma) = \begin{pmatrix} \mathcal{J}_{x,x} & 0 \\ 0 & \mathcal{J}_{\sigma\sigma} \end{pmatrix}$ $\mathcal{J}_{x,x}$ de dimension rxr avec de plus $\mathcal{J}_{xx} = i(\sigma)\mathcal{J}^o_{xx}$ $\mathcal{J}_{\sigma\sigma}$ de dimension $p_2 x p_2$

Alors en posant $\tilde{\mathcal{J}}_{[0,t]}(x,x) = \int_0^t f'(s)\, \mathcal{J}^o_{xx}\, f(s)\, ds$

et en supposant f "orthonormalisée" i.e. $\tilde{\mathcal{J}}^o_{[0,1]}(x,x) = I_{p_1 x p_1}$, on obtient avec normalisation interne (en remarquant qu'on aurait pu aussi faire une combinaison des 2 normalisations interne et externe), la

Proposition (2.6)

Sous les conditions A à F, les tests de région de rejets :

$$\{\sup_k \Lambda_n(k,\ 2[\tilde{\mathcal{J}}_{[0,\frac{k}{n}]}(x,x)]\hat{\eta}^*_{n,k},\ 2\tfrac{k}{n}\hat{\sigma}^*_{n,k}],\ [2[I - \tilde{\mathcal{J}}_{[0,\frac{k}{n}]}(x,x)]\hat{\eta}^{**}_{n,k},\ 2(1-\tfrac{k}{n})\hat{\sigma}^{**}_{n,k}];$$

$$[2\tilde{\mathcal{J}}_{[0,\frac{k}{n}]}(x,x)\hat{\eta}_{n,n},\ \tfrac{2k}{n}\hat{\sigma}_{n,n}]) > A\}$$

et $\{ \operatorname{Sup}_k \Lambda_n(k, \left[2 \eta^{(x,x)}_{[0,\frac{k}{n}]} \hat{\eta}^{*}_{n,k},\ 2\frac{k}{n}\hat{\sigma}^{*}_{n,k}\right], \left[2\left[I - \eta^{(x,x)}_{[0,\frac{k}{n}]}\right]\hat{\eta}^{**}_{n,k},\ 2(1-\frac{k}{n})\hat{\sigma}^{**}_{n,k}\right];$

$\left[2\left[I - \eta^{x,x}_{[0,\frac{k}{n}]}\right]\hat{\eta}_{n,n},\ 2(1-\frac{k}{n})\hat{\sigma}_{n,n}\right]) > A\}$

sont de niveau approximable par :

$$P\{ \operatorname{Sup}_{t\in[0,1]} [\overset{\vee}{X}(t) - \eta^{(x,x)}_{[0,t]}\overset{\vee}{X}(1)]' \left[\overset{\vee}{X}(t) - \eta^{(x,x)}_{[0,t]}\overset{\vee}{X}(1)\right] + \sum_{i=1}^{p_2} B^{(i)}(t)^2 > \operatorname{Log} A\}$$

avec

$$\overset{\vee}{X}(t) = \int_0^t f(s)' \, \eta^{\circ}(x,x)^{\frac{1}{2}} \, dW_r(s).$$

Remarques. Les quantités $\hat{\theta}^{*}_{n,k}$, $\hat{\theta}^{**}_{n,k}$ représentent, à partir du corollaire 2.2 les vrais estimateurs du maximum de vraisemblance des paramètres avant et après la rupture envisagée. Le passage en toute rigueur de la proposition 2.3 au corollaire 2.2 et aux propositions (2.5) (2.5 bis) (2.6), suppose la remarque préliminaire : (avec les notations de la proposition 2.3)

$$\frac{Z_n((\psi(\frac{k}{n}) - I)\,\theta_o\sqrt{n} + \psi(\frac{k}{n})\,\hat{\theta}_{n,k},\ \frac{k}{n})}{Z_n(\psi(\frac{k}{n})\,\hat{\theta}_{n,k},\ \frac{k}{n})} \times \frac{Z_n(\psi(\frac{k}{n})\,\hat{\theta}_{n,n},\ \frac{k}{n})}{Z_n((\psi(\frac{k}{n}) - I)\theta_o\sqrt{n} + \psi(\frac{k}{n})\,\hat{\theta}_{n,n},\ \frac{k}{n})}$$

tend en loi sous $H_{o,n}(\theta_o)$ vers 1 (ce qui est une conséquence facile de la proposition 2.3 et d'un développement de Taylor au 2nd ordre tel qu'il figure dans la démonstration du lemme 4.2, (ainsi que la formule analogue en temps retourné).

III.- ESTIMATION PAR LE MAXIMUM DE VRAISEMBLANCE - LOIS LIMITES

Le problème ici est, en présence d'une rupture d'estimer des trois paramètres $(\overline{\theta}, d, \frac{k}{n})$. La méthode classique (IBRAGIMOV-KHASMINSKII [9]) d'obtention de la loi limite de l'estimateur du maximum de vraisemblance dans un problème standard consiste à paramétrer convenablement un voisinage (au sens de la convergence des expériences (LE CAM [11])) de l'hypothèse $H_{1,n}$ de sorte ensuite à prouver la convergence de la suite des processus de vraisemblance dans un espace fonctionnel qui permet de déduire sans efforts la convergence de la suite des estimateurs de vraisemblance : (exemples : $\mathscr{C}(K)$ si l'ensemble des paramètres est compact, $\mathscr{C}_o$ s'il ne l'est pas). Nous allons adopter ici cette méthodologie à l'aide du théorème d'Invariance de la section 1.

Nous nous restreindrons à estimer une rupture ne se produisant pas en début ou fin d'observation, c'est-à-dire du type $H_{1,n}(t_o,\overline{\theta},d_n)$, $t_o \in]0,1[$, $\overline{\theta} \in \mathbb{R}^p$.

Il reste à définir le comportement de d_n :

Supposons que d_n ne tende pas vers 0 $(d_n = d)$. On montre facilement, sous de faibles conditions que pour qu'une suite d'hypothèses $H_{1,n}(\frac{k_n}{n}, \overline{\theta}_n, \delta_n)$ soit contigue à $H_{1,n}(t_o, \overline{\theta}, d)$ il est nécessaire que $nt_o - k_n$ reste borné. En écrivant le rapport de vraisemblance, on peut alors montrer simplement que dans ce cas, l'estimateur $\hat{k}_n$ du maximum de vraisemblance de $k = [n\, t_o]$ aura une loi limite portée par Z et s'évanouissant très vite quand on s'éloigne de nt_o. Dans ce cas, ce n'est donc pas le passage en temps continu (estimation de $\frac{k}{n}$) qui donne des résultats mais des méthodes de fluctuations. Ce cas est celui résolu par Hinkley [7], [8]. L'intérêt de ce point de vue est de donner une loi

limite directement discrétisée. En revanche, les lois limites obtenues dépendent lourdement de la loi initiale des variables et de l'amplitude de la rupture, ce qui rend nécessaires des calculs d' approximations pour chaque cas particulier.

Une manière de court circuiter ce problème consiste à faire tendre d_n vers 0. Intuitivement, cela diminuera la "concentration" de la loi limite de $\hat{k}_n$ autour de nt_o, puisque l'amplitude de la rupture est moins forte. Ce qui explique que dans ce cas, on se tourne vers l'estimation non plus de k mais d'une fonction de k et de n : (exemple $\frac{k}{n}$ ou $T_n(k)$ au paragraphe 2).

Ce faisant, il est toutefois intéressant de noter que cette condition ($d_n \to 0$) ne suffit pas, à cause des problèmes de bords, à cerner le problème.

Deux cas seront envisagés : 1) $\sqrt{n}\, d_n \longrightarrow d$, les problèmes de bords rencontrés pour le test du rapport de vraisemblance se retrouvent semblables ici et il conviendrait de renormaliser la vraisemblance pour obtenir des lois limites. Cependant il ne s'avèrera pas intéressant de le faire : cela pour deux raisons :

D'abord, cela ne serait pas "statistiquement raisonnable" dans un tel cas on veut estimer une rupture aussi petite que le degré de finesse de l'estimation, ce qui rend impossible l'estimation consistante de l'instant de cette rupture ; la seconde raison est que dans le second cas ($\sqrt{n}\, d_n$ non borné), il est possible de faire une reparamétrisation qui permet l'obtention de lois asymptotiques simples à calculer.

Ce type de renormalisation a été étudiée dans BHATTACHARYA-BROCKWELL [2] sur un modèle gaussien, annoncé pour un modèle correspondant à l'exemple 1, du présent article dans [1], démontré dans [3] ; cependant la méthodologie

choisie n'étant pas la même, l'auteur est obligé de s'intéresser à un estimateur de maximum de vraisemblance modifié, il n'envisage pas le cas d'une régression dépendant du temps, enfin ses restrictions sur le modèle sont très fortes.

1°) Estimation pour une rupture de petite amplitude : $\|\sqrt{n}\, d_n - d_o\| \to 0, \|d_o\| > 0$

Considérons la suite de processus de $\mathcal{C}([0,1] \times \mathbb{R}^{2p})$ définie par interpolation linéaire sur les points :

$$(\frac{k}{n},\bar{\theta},d) \longrightarrow \Lambda_n^1(\frac{k}{n},\bar{\theta},d) = \left[\frac{dP_{H_{1,n}}(\frac{k}{n},\bar{\theta}_o + \frac{\bar{\theta}}{\sqrt{n}},\, d_n + \frac{d}{\sqrt{n}})}{dP_{H_{1,n}}(t_o,\bar{\theta}_o,d_n)}\right]^{1/q}$$

Théorème (3.1)

Sous les conditions de la proposition (1.2) la suite Λ_n^1 converge étroitement, sous $H_{1,n}(t_o,\bar{\theta}_o,d_n)$ dans $\mathcal{C}_\Phi$ vers le processus :

$$\Lambda^1(t,\bar{\theta},d) = \exp \frac{1}{q}\{\, \bar{\theta}'\, \mathcal{I}_{[0,1]}^{-1}\, Y(1) + d'(\mathcal{I}_{[0,t]}\, \mathcal{I}_{[0,1]}^{-1}\, Y(1) - Y(t))$$

$$+ d_o'[(\mathcal{I}_{[0,t]} - \mathcal{I}_{[0,t_o]})\, \mathcal{I}_{[0,1]}^{-1}\, Y(1) - (Y(t) - Y(t_o))]$$

$$- \frac{1}{2}\, \bar{\theta}'\, \mathcal{I}_{[0,1]}\, \bar{\theta} - \frac{1}{2}\, (d+d_o)'\, \mathcal{I}_{[0,t]}(I_{pp} - \mathcal{I}_{[0,1]}^{-1}\, \mathcal{I}_{[0,t]}))(d+d_o)$$

$$- \frac{1}{2}\, d_o'\, \mathcal{I}_{[0,t_o]}\, (I_{p,p} - \mathcal{I}_{[0,t]}^{-1}\, \mathcal{I}_{[0,t_o]})\, d_o$$

(Pour alléger les notations, nous écrivons désormais $\mathcal{I}_{[0,t]}$ au lieu $\mathcal{I}_{[0,t]}(\sigma_o)$).

Démonstration : Le théorème est en fait un corollaire facile de la proposition (1.2).

En effet, il résulte de cette proposition d'abord que les suites de probabilités $P_{H_{1,n}(t_o,\bar{\theta}_o,d_n)}$ et $P_{H_{0,n}(\bar{\theta}_o)}$ sont contigues ce qui permet de déduire immédiatement la relative étroite compacité sous $H_{1,n}(t_o,\bar{\theta}_o,d_n)$ de Λ_n^1 dans $\mathscr{C}_\Phi$ et la convergence des marginales finies s'obtient en appliquant le théorème (7.1) de Roussas [15]. Il est donc bien clair, à la suite de ce théorème, en reprenant les arguments de la section précédente que l'estimateur du maximum de vraisemblance de t a tendance à se concentrer sur les points 0 et 1.

Pour les raisons expliquées dans l'introduction de ce paragraphe, nous ne poursuivrons pas l'étude de ce cas plus avant.

2°) Estimation pour une rupture d'amplitude moyenne : $\|d_n\| \to 0$, $\|\sqrt{n}\, d_n\| \to +\infty$

Considérons le changement de paramétrage de $[0,1]\times\mathbb{R}^{2p}$ dans $\mathbb{R}^{2p+1}$

$(\frac{k}{n},\bar{\theta},d) \longrightarrow (T,\bar{\theta},U)$ avec :

$$T = T_n(k) = n\, d_n' \left(\mathcal{J}_{[0,\frac{k}{n}]} - \mathcal{J}_{[0,t_o]} \right) d_n$$

$$U = U(d) = \mathcal{J}_{[0,t_o]}^{1/2} \left(I_{pp} - \mathcal{J}_{[0,1]}^{-1} \mathcal{J}_{[0,t]}\right)^{1/2} d$$

et la nouvelle suite de processus Λ_n^2 de $\mathfrak{C}(\mathbb{R}^{2p+1})$ définie sur $\mathbb{R}$, à $\bar{\theta}$ et U fixés, par linéarisation sur

les points : $(-n\, d_n'\, \mathcal{J}_{[0,t_o]}\, d_n - 1,\ 0)$,

$$\left(T_n(k),\ \left(\frac{dP_{H_{1,n}}(\frac{k}{n},\ \bar{\theta}_o + \frac{\bar{\theta}}{\sqrt{n}},\ d_n + (I_{pp} - \mathcal{J}_{[0,1]}^{-1}\cdot \mathcal{J}_{[0,t_o]})^{-\frac{1}{2}} \mathcal{J}_{[0,t_o]}^{-\frac{1}{2}} \frac{U}{\sqrt{n}})}{dP_{H_{1,n}}(t_o,\ \bar{\theta}_o,\ d_n)}\right)^{1/q}\right.$$

$k = 1,\ldots,[n\, t_o]$,

$$(T_n(k),\ (\frac{dP_{H_{1,n}}(\frac{k}{n}, \overline{\theta}_o + \frac{\overline{\theta}}{\sqrt{n}},\ d_n + (I_{pp} - \mathcal{J}^{-1}_{[0,1]} \cdot \mathcal{J}_{[0,t_o]})^{-\frac{1}{2}} \mathcal{J}^{-\frac{1}{2}}_{[0,t_o]} \frac{U}{\sqrt{n}})^{\frac{1}{q}}}{d\,P_{H_{1,n}}(t_o, \overline{\theta}_o, d_n)}$$

$$k = [n\,t_o] + 1, \ldots, n \quad (n\ d_n'\ \mathcal{J}_{[t_o,1]}\ d_n + 1,\ 0)$$

et prolongé par 0 en dehors de l'intervalle

$$[-n\ d_n'\ \mathcal{J}_{[0,t_o]} d_n - 1\ ,\ n\ d_n'\ \mathcal{J}_{[t_o,1]}\ d_n + 1] \ .$$

Théorème (3.2)

Supposons vérifiées les hypothèses A à H , s'il existe $\delta \in]0, \frac{1}{2}[$ tel que $n^{\frac{1}{2}-\delta} \| \mathcal{J}^{1/2}\ \mathcal{F}(t_o) d_n \| \longrightarrow +\infty$, alors sous $H_{1,n}(t_o, \overline{\theta}_o, d_n)$:

1) La suite de processus Λ^1_n converge vers 0 uniformément sur les compacts de $([0,1] \setminus \{t_o\} \times \mathbb{R}^{2p}$ en probabilité .

2) La suite de processus Λ^2_n converge étroitement dans $\mathcal{C}_o(\mathbb{R}^{2p+1})$ vers le processus :

$$\Lambda^2(T,\overline{\theta},U) = \exp \frac{1}{q}\{(\ \overline{\theta}'\ \mathcal{J}^{\frac{1}{2}}_{[0,1]}\ \xi_1 + U'\ \mathcal{J}^{\frac{1}{2}}_{[0,1]}\ \xi_2 + Z(T))$$

$$- \frac{1}{2}\ (\ \overline{\theta}'\ \mathcal{J}_{[0,1]}\ \overline{\theta} + U'\ \mathcal{J}_{[0,1]}\ U + |T|)\}$$

où ξ_1 et ξ_2 sont des variables normales $N(0, I_{p \times p})$, indépendantes, $Z(T)$ est un mouvement brownien unidimensionnel sur $\mathbb{R}$ passant par 0 en 0 et indépendant de ξ_1 et ξ_2 .

Démonstration : Nous reprenons les notations précédantes :

$$\theta^*_n = \mathcal{J}^{-1}_{[0,1]} \overline{\theta} - \mathcal{J}^{-1}_{[0,1]}\ \mathcal{J}_{[t_o,1]} d_n,\ \theta^{**}_n = \mathcal{J}^{-1}_{[0,1]} \overline{\theta} + \mathcal{J}^{-1}_{[0,1]}\ \mathcal{J}_{[0,t_o]} d_n \ , \text{ et}$$

nous posons :

$$\theta^* = \mathcal{I}^{-1}_{[0,1]}\bar{\theta} - \mathcal{I}^{-1}_{[0,1]}\,\mathcal{I}_{[\frac{k}{n},1]}\,d + \mathcal{I}^{-1}_{[0,1]}\left[\mathcal{I}_{[0,\frac{k}{n}]} - \mathcal{I}_{[0,t_o]}\right]\sqrt{n}\,d_n$$

$$\theta^{**} = \mathcal{I}^{-1}_{[0,1]}\bar{\theta} + \mathcal{I}^{-1}_{[0,1]}\,\mathcal{I}_{[0,\frac{k}{n}]}\,d + \mathcal{I}^{-1}_{[0,1]}\left[\mathcal{I}_{[0,\frac{k}{n}]} - \mathcal{I}_{[0,t_o]}\right]\sqrt{n}\,d_n$$

En loi, nous pouvons écrire l'égalité :

$$\Lambda^1_n(\tfrac{k}{n}, \bar{\theta}, d) = \Lambda^2_n(T_n(k), \bar{\theta}, U(d))$$

$$= Z^{\cdot}_{n,\theta^*_n}(\theta^*, \tfrac{k}{n})\,\frac{Z^{\cdot}_{n,\theta^*_n}(\theta^{**} + \sqrt{n}\,d_n,\ t_o)}{Z^{\cdot}_{n,\theta^*_n}(\theta^{**} + \sqrt{n}\,d_n, \tfrac{k}{n})}$$

$$\times\ \frac{Z^{\cdot}_{n,\theta^{**}_n}(\theta^{**}, 1)}{Z^{\cdot}_{n,\theta^{**}_n}(\theta^{**}, t_o)}\ 1_{\{\frac{k}{n} \leq t_o\}}$$

$$+ Z^{\cdot}_{n,\theta^*_n}(\theta^*, t_o) \times \frac{Z^{\cdot}_{n,\theta^{**}_n}(\theta^{**} - \sqrt{n}\,d_n, \tfrac{k}{n})}{Z^{\cdot}_{n,\theta^{**}_n}(\theta^{**} - \sqrt{n}\,d_n, t_o)} \times \frac{Z^{\cdot}_{n,\theta^{**}_n}(\theta^{**}, 1)}{Z^{\cdot}_{n,\theta^{**}_n}(\theta^{**}, \tfrac{k}{n})}\ 1_{\{\frac{k}{n} > t_o\}}$$

la suite de processus à 2 dimensions $(Z^{\cdot}_{n,\theta^*_n}, Z^{\cdot}_{n,\theta^{**}_n})$ étant obtenue

à partir d'un n échantillon du 2-vecteur $(X^{\cdot}_{\theta^*_n}, Y^{\cdot}_{\theta^{**}_n})$ de loi

$P_{\theta^*_n} \otimes P_{\theta^{**}_n}$ en construisant $Z^{\cdot}_{n,\theta^*_n}$ (resp. $Z^{\cdot}_{n,\theta^{**}_n}$) à partir des

$X^i_{\theta^*_n}$ (resp. $Y^i_{\theta^{**}_n}$) . Les 2 termes de la somme se comportant symétriquement

dans tous les problèmes traités dans la suite, systématiquement, nous

n'envisagerons que le cas $\frac{k}{n} \leq t_o$ $(T_n(k) \leq 0)$.

La convergence des marginales finies étant de routine, nous concentrerons notre attention sur les inégalités suivantes qui assurent la tension sur les compacts et le contrôle à l'infini :

$\forall \varepsilon > 0$, $\forall T_1 \leq 0$, $\forall M > 0$, $\exists\ \delta$ tel que :

i) $\mathbb{P}_1 = P\Big(\sup_{\substack{T,T': \\ T_1 \leq T-\delta \leq T' \leq T \leq 0}} \ \sup_{\|(\bar\theta,U)\| < M} |\Lambda_n^2(T,\bar\theta,U) - \Lambda_n^2(T',\bar\theta,U)| > \varepsilon \Big) \leq \varepsilon$

ii) $\mathbb{P}_2 = P\Big(\sup_{\substack{T: \\ T_1 \leq T \leq 0}} \ \sup_{\substack{\|(\bar\theta,U)-(\bar\theta^1,U^1)\| \leq \delta \\ \|(\bar\theta,U)\| \leq M}} |\Lambda_n^2(T,\bar\theta,U) - \Lambda_n^2(T,\bar\theta^1,U^1)| > \varepsilon \Big) \leq \varepsilon$

$\forall \varepsilon > 0$, $\exists\ T_1 \leq 0$ tel que :

iii) $\mathbb{P}_3 = P\Big(\sup_{T \leq T_1} \ \sup_{(\bar\theta,U) \in \mathbb{R}^{2p}} \Lambda_n^2(T,\bar\theta,U) > \varepsilon \Big) \leq \varepsilon$

$\forall \varepsilon > 0$, $\forall T_1 \leq 0$, $\exists M > 0$ tel que :

iv) $\mathbb{P}_4 = P\Big(\sup_{T_1 \leq T \leq 0} \ \sup_{\|(\bar\theta,U)\| > M} \Lambda_n^2(T,\bar\theta,U) > \varepsilon \Big) \leq \varepsilon$

Remarque : L'inégalité (iii) prouvera en même temps l'assertion 1) du théorème.

Pour montrer les inégalités (i) et (ii) qui assurent la tension sur les compacts, il est commode de remarquer que $\Lambda_n^2(T_n(k), \bar\theta, U(d))$ se divise en deux parties :

$$Z_{n,\theta_n^*}(\theta^*, \tfrac{k}{n}) \times \frac{Z_{n,\theta_n^{**}}(\theta^{**}, 1)}{Z_{n,\theta_n^{**}}(\theta^{**}, t_o)} \quad \text{et} \quad \frac{Z_{n,\theta_n^*}(\theta^{**} + \sqrt{n}\, d_n,\ t_o)}{Z_{n,\theta_n^*}(\theta^{**} + \sqrt{n}\, d_n,\ \frac{k}{n})} \quad \text{qui séparent en}$$

fait les paramètres en $(\overline{\theta}, U(d))$ d'une part et $T_n(k)$ d'autre part, comme le montre les 2 lemmes suivants :

Lemme (3.1)

Pour tout K compact de $\mathbb{R}^p$, pour tout $T_1 < 0$. Pour tout ε positif il existe n_o tel que :

$$\forall\, n > n_o \quad , \quad P\Big(\sup_{\theta\in K} \sup_{k\in K_n(T_1)} \left| \frac{Z_{n,\theta_n^*}(\sqrt{n}\, d_n,\, t_o)}{Z_{n,\theta_n^*}(\sqrt{n}\, d_n,\, \frac{k}{n})} - \frac{Z_{n,\theta_n^*}(\theta + \sqrt{n}\, d_n, t_o)}{Z_{n,\theta_n^*}(\theta + \sqrt{n}\, d_n, \frac{k}{n})} \right| > \varepsilon \Big) \leq \varepsilon$$

$K_n(T_1) = \{k,\ T_1 \leq T_n(k) \leq 0\}$.

Démonstration

Pour alléger les notations, remarquons que :

$$\frac{Z_{n,\theta_n^*}(\theta, t_o)}{Z_{n,\theta_n^*}(\theta, \frac{k}{n})} = \tilde{Z}_n^{\tilde{f}}(\theta,\ t_o - \frac{k}{n})$$

où $\tilde{f}(t) = f(t_o - t)$ et $\tilde{Z}_n$ est le processus construit sur les variables en ordre renversé : $Y_{\theta_n^*, i} = X_{\theta_n^*}^{nt_o - i - 1}$, l'indice θ_n^* sera désormais omis sans prêter à confusion. Soit $k_o = \dfrac{|T_1|}{\varepsilon^2}$

$$P\Big(\sup_{\theta\in K} \sup_{k\in K_n(T_1)} \left| \tilde{Z}_n^{\tilde{f}}(\sqrt{n}\, d_n,\, t_o - \tfrac{k}{n}) - \tilde{Z}_n^{\tilde{f}}(\theta + \sqrt{n}\, d_n,\, t_o - \tfrac{k}{n}) \right| > \varepsilon\Big)$$

$$\leq P\Big(\sup_{\theta\in K} \sup_{nt_o - k < k_o} \left| \tilde{Z}_n^{\tilde{f}}(\sqrt{n}\, d_n,\, t_o - \tfrac{k}{n}) - \tilde{Z}_n^{\tilde{f}}(\theta + \sqrt{n}\, d_n,\, t_o - \tfrac{k}{n}) \right| > \varepsilon\Big)$$

$$+ P\Big(\sup_{\theta\in K} \sup_{k\in K_n(T_1,k_o)} \Big|\tilde{Z}_n^{\tilde{f}}(\sqrt{n}\,d_n,\ t_o-\tfrac{k}{n}) - \tilde{Z}_n^{\tilde{f}}(\theta+\sqrt{n}\,d_n,\ t_o-\tfrac{k}{n})\Big| > \varepsilon\Big) \qquad (3.1)$$

$$K_n(T_1,k_o) = K_n(T_1) \cap \{nt_o - k \geq k_o\} .$$

Etudions d'abord le premier terme de (3.1) :

$$\text{si } d_n = (\eta_n', \sigma_n')', \ \forall \ell < \infty,\ P\Big\{\sup_{\theta\in K} \sup_{nt_o-k<\ell} \Big|\tilde{Z}_n^{\tilde{f}}(\sqrt{n}\,d_n, t_o-\tfrac{k}{n}) - \tilde{Z}_n^{\tilde{f}}(\theta+\sqrt{n}\,d_n,\ t_o-\tfrac{k}{n})\Big| > \varepsilon\Big)$$

$$\leq \ell\ P\Big(\sup_{\theta\in K} \sup_{nt_o-i<\ell} \Big|\frac{g(X_i - f(\frac{i}{n})\eta_n,\ \frac{\sigma}{\sqrt{n}}+\sigma_o)}{g(x_i,\ \sigma_o)} - \frac{g(X_i - f(\frac{i}{n})(\frac{\eta}{\sqrt{n}}+\eta_n),\ \frac{\sigma}{\sqrt{n}}+\sigma_n+\sigma_o)}{g(x_i,\ \sigma_o)}\Big| > \gamma_A(\varepsilon)\Big)$$

$$+ \ell\ P\Big(\sup_{\theta\in K, nt_o-i<\ell} \Big|\frac{g(X_i - f(\frac{i}{n})\eta_n\ ,\ \frac{\sigma}{\sqrt{n}}+\sigma_o)}{g(x_i,\ \sigma_o)}\Big| > A\Big) \qquad (3.2)$$

où $\gamma_A(\varepsilon) = \dfrac{\varepsilon}{\ell\ A^{\ell}}$.

Or les hypothèses A(ii) et H impliquent l'existence de A et de n_o tels que (3.2) soit inférieur à $\frac{\varepsilon}{3}$ pour $n > n_o$ (compte tenu du fait que $\|d_n\| \to 0$).

Le second terme de la somme (3.1) se majore à l'aide d'un développement de Taylor par :

$$P\Big(\sup_{k\in K_n(T_1,k_o)} d_n' \sum_{i=1}^{nt_o-k} \dot{\tilde{\mathscr{F}}}(\tfrac{i}{n}) L(Y_i,\sigma_o) + Q_n^2(\sqrt{n}\,d_n, \tfrac{k}{n}) > \log A\Big) +$$

$$P\Big(\sup_{\theta\in K}\sup_{k\in K_n(T_1,k_o)}\Big|\frac{\theta'}{\sqrt{n}}\sum_{i=1}^{nt_o-k}\tilde{\mathcal{F}}\,(\tfrac{i}{n})'\dot{L}(Y_i,\sigma_o)+Q_n^1(\theta+\sqrt{n}\,d_n,\tfrac{k}{n})$$

$$-Q_n^2(\sqrt{n}\,d_n,\tfrac{k}{n})\Big|>\frac{\varepsilon}{A}\Big)$$

avec $Q_n^\nu(\theta,k)=\theta'\frac{1}{2n}\sum_{i=1}^{nt_o-k}\tilde{\mathcal{F}}(\tfrac{i}{n})'\ddot{L}\,(Y_i-f(\tfrac{i}{n})\,\frac{\eta_k^\nu(n)}{\sqrt{n}}\,,\,\sigma_o+\frac{\sigma_k^\nu(n)}{\sqrt{n}}\,)\,\tilde{\mathcal{F}}(\tfrac{i}{n})\,\theta$

pour $\nu=1$ et 2,

et $\theta_k^1(n)\in[0,\sqrt{n}d_n+\theta]$, $\theta_k^2(n)\in[0,\,d_n\sqrt{n}]$

soit encore par :

$$P\Big(\sup_{\theta\in K}\sup_{k\in K_n(T_1,k_o)}\Big|\frac{\theta'}{\sqrt{n}}\sum_{i=1}^{nt_o-k}\tilde{\mathcal{F}}\,(\tfrac{i}{n})'\dot{L}(Y_i,\sigma_o)\Big|>\frac{\varepsilon}{3A}\Big)+$$

$$P\Big(\sup_{k\in K_n(T_1,k_o)}d_n'\sum_{i=1}^{nt_o-k}\tilde{\mathcal{F}}\,(\tfrac{i}{n})'\dot{L}(Y_i,\sigma_o)>\frac{\log A}{2}\Big)$$

$$+P\Big(\sup_{k\in K_n(T_1,k_o)}Q_n^2(\sqrt{n}\,d_n,\tfrac{k}{n})>\frac{\log A}{2}\Big)$$

$$+P\Big(\sup_{\substack{\theta\in K\\ k\in K_n(T_1,k_o)}}\Big|Q_n^1(\theta+\sqrt{n}\,d_n,\tfrac{k}{n})-Q_n^1(\sqrt{n}\,d_n,\tfrac{k}{n})\Big|>\frac{\varepsilon}{3A}\Big)$$

$$+P\Big(\sup_{k\in K_n(T_1,k_o)}\Big|Q_n^1(\sqrt{n}\,d_n,\tfrac{k}{n})-Q_n^2(\sqrt{n}\,d_n,\tfrac{k}{n})\Big|>\frac{\varepsilon}{3A}\Big)=J\ .$$

L'inégalité maximale (cf. par exemple [22] théorème 20) permet de majorer les 2 premiers termes de la somme et on obtient pour $k>n_o$, $k_n(T)$ déterminé par la relation $T_n(k_n(T))=T$, et $|K|=\sup_{\theta\in K}|\theta|$

$$J<\frac{9\|J_{[0,t_o]}-J_{[0,k_n(T_1)]}\|\,|K|A^2}{\varepsilon^2}+\frac{T_1^2}{(\log A)^2}+P\Big(\sup_{k\in K_n(T_1,k_o)}d_n'\sum_{i=1}^{nt_o-k}\tilde{\mathcal{F}}\,(\tfrac{i}{n})'\ldots$$

$$\tilde{\mathcal{F}}(\tfrac{i}{n})\ d_n\ \bar{\lambda}(Y_i,\delta_1) > \frac{\log A}{2}$$

$$+ P\Big(\sup_{k\in K_n(T_1,k_o)} d_n' \sum_{i=1}^{nt_o-k} \tilde{\mathcal{F}}(\tfrac{i}{n})' \tilde{\mathcal{F}}(\tfrac{i}{n})\ [\bar{\lambda}(Y_i,\delta_1) - \underline{\lambda}(Y_i,\delta_1)]\, d_n > \tfrac{\varepsilon}{3}\Big)$$

$$+ P\Big(\sup_{\substack{\theta\in K \\ k\in K_n(T_1,k_o)}} |\theta'\ \frac{1}{n} \sum_{i=1}^{nt_o-k} \tilde{\mathcal{F}}(\tfrac{i}{n})' \tilde{\mathcal{F}}(\tfrac{i}{n})\ \bar{\lambda}(Y_i,\delta_1)\ \theta| > \tfrac{\varepsilon}{3}\Big) \qquad (3.3)$$

où $\bar{\lambda}(Y_i,\delta_1)$ et $\underline{\lambda}(Y_i,\delta_1)$ désignent respectivement le suprémum de la plus grande valeur propre et l'infimum de la plus petite valeur propre de la matrice $\ddot{L}(Y_i-\mu,\ \sigma_o+\sigma)$ sur l'ensemble $\{\|(\mu,\sigma)\| < \delta_1\}$.

En reprenant les notations du lemme 4.7. qui permet maintenant de majorer les 3 derniers termes par ε pour $n > n_o$ (en réglant ε' à $\frac{\varepsilon}{6T_1A}$) compte tenu du fait que les inégalités suivantes sont toujours vérifiées:

$$\sup_{k\in K_n(T_1,n_o)} 2\,\frac{\varepsilon}{6T_1A}\ n\ d_n'\Big(\mathcal{J}_{[0,t_o]} - \mathcal{J}_{[0,\frac{k}{n}]}\Big)d_n < \frac{\varepsilon}{A}$$

$$\sup_{\substack{\theta\in K \\ k\in K_n(T_1,n_o)}} \theta'\Big(\mathcal{J}_{[0,t_o]} - \mathcal{J}_{[0,\frac{k}{n}]}\Big)\theta < \varepsilon$$

La démonstration s'achève en majorant le premier terme de (3.3) par

$$\frac{C|T_1|}{\varepsilon^2\,\|\sqrt{n}\ \mathcal{F}(t_o)\ \mathcal{J}(\sigma_o)d_n\|^2}$$

Lemme (3.2)

Pour tout K compact de $\mathbb{R}^{2p}$, pour tout $T_1 < 0$, pour tout ε positif il existe n_o tel que pour $n > n_o$, on a :

$$P\Big(\sup_{\theta\in K}\ \sup_{k\in K_n(T_1)} |Z^{\cdot}_{n,\theta^*_n}(\theta,\tfrac{k}{n}) - Z^{\cdot}_{n\theta^*_n}(\theta,t_o)| > \varepsilon\Big)\ .$$

Ce lemme est un corollaire trivial du théorème 1 compte tenu du fait que

si $k \in K_n(T_1)$, pour tout δ fixé il existe n_o tel que si $n > n_o$ alors :

$$|\frac{k}{n} - t_o| < \delta$$

Les inégalités (i), (ii) sont alors un corollaire des 2 lemmes précédents et de la convergence de la suite de processus $\{\overset{\vee}{Z}_n^f(T),\ T \in \bar{R}\}$: défini par linéarisation sur les points :

$$\{(-n\, d_n'\ \mathfrak{I}_{[0,t_o]} d_n - 1, 0),\ (T_n(k), \tilde{Z}_n(\sqrt{n}\, d_n, t_o - \frac{k}{n}))\ k=1,\ldots,[nt_o], (0,1)\}$$

et prolongé par 0 au delà de $-n\, d_n'\ \mathfrak{I}_{[0,t_o]}\ d_n - 1$.

Lemme (3.3)

La suite $\{\overset{\vee}{Z}_n(T)\ ,\ T \in \mathbb{R}^-\}$ converge étroitement sur $\mathscr{C}(\mathbb{R}^-)$ muni de la topologie de la convergence uniforme sur les compacts vers $Z(T) = \exp \frac{1}{q}\{W(T) + \frac{T}{2}\}$.

Démonstration

On a vu que pour $nt_o - k > k_o$, $\log \overset{\vee}{Z}_n(T_n(k))$ peut être développé en une somme de 2 termes :

$$d_n' \sum_{i=1}^{k} \overset{\sim}{\mathscr{F}}(\tfrac{i}{n}) L(Y_i, \sigma_o) + \frac{1}{2}\, d_n' \sum_{i=1}^{k} \overset{\sim}{\mathscr{F}}(\tfrac{i}{n}) L(Y_i - f(\tfrac{i}{n}) \frac{\eta_k^*(n)}{\sqrt{n}}\ ,$$

$$\sigma_o + \frac{\sigma_k(n)}{\sqrt{n}}\ )\ \overset{\sim}{\mathscr{F}}(\tfrac{i}{n})\ d_n$$

avec $\theta_k^*(n) \in [0, \sqrt{n}\, d_n]$.

Le premier terme converge, c'est une conséquence, par exemple de Rebolledo ([14], corollaire 14, p. 86).

La loi des grands nombres énoncée dans le lemme 4.7. permet de montrer la convergence du second pour $k > n_o$.

Il reste à traiter le cas des n_o premiers k : $nt_o - k < n_o$ pour lesquels pour tout δ^* positif il existe n'_o tel que $n > n'_o$ implique $T_n(k) < \delta^*$:

$$P\Big(\sup_{nt_o - k < n_o} \Big| \prod_{i=1}^{k} \frac{g(X_i - f(\tfrac{i}{n})\eta_n\ ,\ \sigma_o + \sigma)}{g(X_i,\ \sigma_o)} - 1 \Big| > \varepsilon \Big)$$

$$< n_o^2\, P\Big(\sup_{|(\mu,\sigma)|<\delta^*} \Big| \frac{g(X_i - \mu,\ \sigma_o + \sigma)}{g(X_i, \sigma_o)} - 1 \Big| > \gamma_1(\varepsilon) \Big)$$

en reprenant les notations de la formule (3.2) et en achevant de même .

Montrons maintenant l'inégalité (iv). Pour cela, nous majorons la probabilité $\mathbb{P}_4$ par la somme : $\pi_1 + \pi_2$ où

$$\pi_1 = P\Big(\sup_{(k,\bar{\theta},U)\in B_1} \Lambda_n^2(T_n(k),\bar{\theta},U) > \varepsilon \Big)$$

$$\pi_2 = P\Big(\sup_{(k,\bar{\theta},U)\in B_2} \Lambda_n^2(T_n(k),\bar{\theta},U) > \varepsilon \Big)$$

$$B_1 = \{(k,\bar{\theta},U)/T_1 \leq T_n(k) \leq 0\ ;\ ||\bar{\theta},\ U|| > M\ ;\ ||\theta^{**} + \sqrt{n}\, d_n|| \leq 2||\sqrt{n}\ d_n||\}$$

$$B_2 = \{(k,\bar{\theta},U)/T_1 \leq T_n(k) \leq 0\ ;\ ||\theta^{**} + \sqrt{n}\, d_n|| > 2\ ||\sqrt{n}\ d_n||\}$$

La majoration de π_1 consiste en 2 étapes :

1) Montrer que sur B_1, $\tilde{Z}_n^f(\theta^{**} + \sqrt{n}\, d_n,\ t_o - \frac{k}{n})$ est borné

2) Remarquer que sur B_1 soit $||\theta^*||$, soit $||\theta^{**}||$ est supérieur à M ce qui, dans le premier cas grâce au lemme 3.5., directement à l'aide du théorème 1 dans le second, assure une majoration de π_1 .

Lemme (3.4)

Pour tout ε positif il existe $A > 0$ et n_o tel que $n > n_o$ implique :

$$P(\sup_{B_1} \tilde{Z}_n^f(\theta^{**} + \sqrt{n}\, d_n,\ t_o - \tfrac{k}{n}) > A) < \varepsilon$$

Démonstration

Le cadre de démonstration est le même que celui du lemme (3.1) : en effet, le choix de B_1 permet de régler la majoration par un développement de Taylor, les arguments qui s'imposent ensuite étant du type loi des grands nombres et inégalité maximale.

Pour les valeurs de $nt_o - k$ plus grandes que k_o, on écrit : $\log \tilde{Z}_n^f(\theta^{**} + \sqrt{n}\, d_n,\ t_o - \frac{k}{n})$ sous la forme :

$$(\frac{\theta^{**}}{\sqrt{n}} + d_n)' \sum_{i=1}^{nt_o-k} \dot{\mathcal{J}}(\tfrac{i}{n}) \dot{L}(Y_i, \sigma_o) + Q_n^1(\theta^{**} + \sqrt{n}\, d_n, \tfrac{k}{n})$$

où l'on écrit Q_n^1 par analogie avec la démonstration du lemme 3.1, et imposant $\theta_k^*(n) \in [0, \theta^{**} + \sqrt{n}\, d_n]$.

Le terme $\sup_{B_1} Q_n^1(\theta^{**} + \sqrt{n}\, d_n, \frac{k}{n})$ est traité comme on l'a fait précédemment pour $k \geqslant k_o$, et on a :

$$P(\sup_{B_1} (\frac{\theta^{**} + \sqrt{n}\, d_n}{\sqrt{n}})' \sum_{i=1}^{nt_o-k} \dot{\mathcal{J}}(\tfrac{i}{n}) \dot{L}(Y_i, \sigma_o) > \log A)$$

$$< P(\sup_{k/T_1 \leq T_n(k) \leq 0} 2\, \|\sqrt{n}\, d_n\|\, \| \sum_{i=1}^{nt_o-k} \frac{\dot{L}(Y_i, \sigma_o)}{\sqrt{n}} \| > \log A)$$

$$< (\frac{1}{\log A})^2\, 4\, \|\sqrt{n}\, d_n\|^2\, \| \mathcal{J}_{[0,t_o]} - \mathcal{J}_{[0, \frac{k_n(T_1)}{n}]} \| \text{ d'après l'inégalité maximale([22])}$$

$$< \left(\frac{1}{\log A}\right)^2 8 \|\sqrt{n}\, d_n\|^2 \left(t_o - \frac{k_n(T_1)}{n}\right) \mathcal{F}(t_o)' \mathcal{J}\mathcal{F}(t_o)$$

$$< \left(\frac{1}{\log A}\right)^2 8\, T_1 \times \frac{\|\sqrt{n}\, d_n\|^2}{\|\sqrt{n}\, \mathcal{F}(t_o)\, \mathcal{J}(\sigma_o) d_n\|^2} \times \mathcal{F}(t_o)' \mathcal{J}\mathcal{F}(t_o)$$

Le traitement des k_o premières valeurs de $nt_o - k$ est de routine.

Lemme (3.5)

Pour tout $\varepsilon > 0$, pour tout $A > 0$ il existe M, n_o tels que $n > n_o$ implique

$$P\Big(\sup_{B_1(M) \cap \|\theta^*\| > M} Z_n(\theta^*, \frac{k}{n}) > \frac{\varepsilon}{A} \Big) < \varepsilon$$

La démonstration repose sur l'inégalité :

$$P\Big(\sup_{B_1(M) \cap \|\theta^*\| > M} Z_n(\theta^*, \frac{k}{n}) > \frac{\varepsilon}{A} \Big) \leq P\Big(\sup_{B_1(M) \cap \|\theta^*\| > M} \frac{Z_n(\theta^*, \frac{k}{n})}{Z_n(\theta^*, \frac{t_o}{2})} > A \Big)$$

$$+ P\Big(\sup_{\|\theta^*\| > M} Z_n(\theta^*, \frac{t_o}{2}) > \frac{\varepsilon}{A^2} \Big)$$

et l'utilisation du théorème 1 en remarquant que $T_1 \leq T_n(k) \leq 0$ implique $k \geq \frac{3t_o}{4}$ pour $n > n_o$.

La majoration de π_2 est une conséquence de la proposition 1.1. et du lemme (4.5) en remarquant que $\|\theta^{**} + \sqrt{n}\, d_n\| > 2\|\sqrt{n}\, d_n\|$ implique $\|\theta^{**}\| > \|\sqrt{n}\, d_n\|$: $\exists$ n_o, A_n (proposition 1.1)

$$P\Big(\sup_{k,\theta} \tilde{Z}_n^{\tilde{f}}(\theta, \frac{nt_o - k}{n}) > A_n\Big) < \varepsilon \qquad n > n_o \qquad \text{et}$$

$$P\Big(\sup_{\|\theta^{**}\| > \|\sqrt{n}\, d_n\|} \frac{Z_n(\theta^{**}, 1)}{Z_n(\theta^{**}, t_o)} > \frac{\varepsilon}{A_n} \Big) < \frac{C\, A_n^{\gamma}}{\|\sqrt{n}\, d_n\|^{N}} \quad \text{(lemme (4.5))}$$

Le terme $Z_n(\theta^*, \frac{k}{n})$ étant borné en probabilité par un argument de même type que le lemme 3.5, un choix de N réalisant $\frac{A_n^\gamma}{\|\sqrt{n}\, d_n\|^N} \longrightarrow 0$ achève la démonstration de (iv). Passons maintenant à celle de (iii) au moyen du lemme suivant :

Lemme (3.6)

Pour tout $\varepsilon > 0$, $\delta > 0$ il existe T_o tel que :

(1) $P\Big(\sup_{(\theta,k)\in C(n,k_n(-T_o))} \tilde{\tilde{Z}}_n^f(\theta + \sqrt{n}\, d_n, \frac{nt_o - k}{n}) > \delta \Big) < \varepsilon$

(2) $P\Big(\sup_{(\theta,k)\in C(n,nt_1)} \tilde{\tilde{Z}}_n^f(\theta + \sqrt{n}\, d_n, \frac{nt_o - k}{n}) > n^{-N} \Big) < \varepsilon \quad \forall N > 0$

dès que $0 < a < \alpha$ et $t_1 < t_o$,

où $C(n,K_n) = \{(\theta,k) / \|\theta\| < n^a , \ k \leqslant K_n\}$ et

$k_n(-T_o)$ défini par $T_n(k_n(-T_o)) = -T_o$.

Démonstration

La démonstration des 2 inégalités est à peu de chose près la même. Nous la faisons simultanément en désignant $C(n,k_n(-T_o))$ par c_1, $C(n,nt_1)$ par c_2 et δ par d_1 , n^{-N} par d_2. Ces deux ensembles ont encore été choisis pour que le problème se résolve à l'aide d'un développement de Taylor et donc des arguments du type de ceux des lemmes (3.1) ou (3.4) en remarquant cependant qu'ici l'inégalité maximale est à remplacer par une inégalité plus fine dérivée de Birn Baum Marshall [5] : en effet, on a :

(1) Pour $n \geqslant n_o$ $c_1 \subset c_2$

(2) $\forall T_o$, $\forall K$ $c_i \cap \{k \leqslant K\} = \emptyset$ pour $n \geqslant n_o$

(3) $\{\theta / (\theta + \sqrt{n}\, d_n)'(\mathcal{J}_{[0,t_o]} - \mathcal{J}_{[0,\frac{k}{n}]})(\theta + \sqrt{n}\, d_n) > (k-nt_o)\delta^2\} = \emptyset$

pour $n \geq (\dfrac{||\mathcal{F}(t_o)||}{\delta})^{2(1-2\alpha)}$ d'après les hypothèses 1) et 2) du théorème (3.2).

Ce qui permet d'écrire : pour $\nu = 1$ ou 2

$$P(\sup_{(\theta,k)\,\epsilon\, c_\nu} \tilde{Z}_n^{\tilde{f}}(\theta + \sqrt{n}\, d_n, \frac{nt_o-k}{n}) > d_\nu) \leq$$

$$P(\sup_{(\theta,k)\epsilon c_\nu} \frac{(\theta + \sqrt{n}\, d_n)'}{\sqrt{n}} \sum_{i=1}^{nt_o-k} \tilde{\mathcal{F}}(\tfrac{i}{n})\dot{L}(Y_i,\sigma_o) - Q_n^2(\theta + \sqrt{n}\, d_n) > \log d_\nu) \leq$$

$$\varepsilon + P(\sup_{(\theta,k)\epsilon c_\nu} \frac{(\theta + \sqrt{n}\, d_n)'}{\sqrt{n}} \sum_{i=1}^{nt_o-k} \tilde{\mathcal{F}}(\tfrac{i}{n})\dot{L}(Y_i,\sigma_o) - Q_\varepsilon(\theta + \sqrt{n}\, d_n) > \log d_\nu)$$

par un raisonnement désormais de routine avec :

$$Q_\varepsilon(\theta) = \frac{1-\varepsilon}{2}\, \alpha\, \theta'(\mathcal{J}_{[0,t_o]} - \mathcal{J}_{[0,\frac{k}{n}]})\, \theta .$$

En remarquant que pour n assez grand, on a : $\log d_\nu + \frac{1}{2} Q_\varepsilon(\theta + \sqrt{n}\, d_n) \geq 0$:

$$P(\sup_{(\theta,k)\epsilon c_\nu} \tilde{Z}_n^{\tilde{f}}(\theta + \sqrt{n}\, d_n, \frac{nt_o-k}{n}) > d_\nu)$$

$$\leq \varepsilon + P(\sup_{(\theta,k)\epsilon c_\nu} \frac{Q_\varepsilon^{*1/2}(\frac{1}{\sqrt{n}} \sum_{i=1}^{nt_o-k} \tilde{\mathcal{F}}(\tfrac{i}{n})\dot{L}(Y_i,\sigma_o))}{Q_\varepsilon^{1/2}(\theta + \sqrt{n}\, d_n)} > \frac{1}{2})$$

avec $Q_\varepsilon^*(\theta) = \dfrac{1-\varepsilon}{2}\, \dfrac{1}{\alpha}\, \theta'(\mathcal{J}_{[0,t_o]} - \mathcal{J}_{[0,\frac{k}{n}]})^{-1}\, \theta$

$$< \varepsilon + P(\sup_{(\theta,k)\epsilon c_\nu} \frac{Q_\varepsilon^{*1/2}(\frac{1}{\sqrt{n}}(\sum_{i=1}^{nt_o-k} \tilde{\mathcal{F}}(\tfrac{i}{n})\dot{L}(Y_i,\sigma_o))}{Q_\varepsilon^{1/2}(\sqrt{n}\, d_n)} > \frac{1}{4})$$

On peut maintenant terminer la démonstration de ce lemme en utilisant l'inégalité de Birnbaum-Marshall d'après le lemme (4.8). Comme dans la démonstration du lemme (4.6) en remplaçant dans ce cas, dans $A_2(\frac{k}{n})$, $h(t)$ par $\frac{c}{n^{2\alpha}} \times 1_{\{t > \frac{T_o}{n^{2\alpha}}\}}$ et en remarquant qu'alors $\int_o^1 h(t)dt \leqslant \frac{c}{T_o}$.

Pour montrer (iii), il suffit alors de décomposer l'espace en zônes :

$$B_{1,n} = \{(\bar{\theta}, d, \tfrac{k}{n}) / |\theta^{**}| \leqslant n^a \ , \ nt_1 \leqslant k \leqslant k_n(-T_o)\} \ ,$$

$$B_{2,n} = \{(\bar{\theta}, d, \tfrac{k}{n}) / |\theta^{**}| \leqslant n^a \ , \ 0 \leqslant k \leqslant nt_1\}$$

$$B_{3,n} = \{(\bar{\theta}, d, \tfrac{k}{n}) / |\theta^{**}| > n^a \ , \ 0 \leqslant k \leqslant k_n(-T_o)\}$$

Sur $B_{1,n}$: $Z_{n,\theta_n^*}(\theta^*, \frac{k}{n}) \dfrac{Z_{n,\theta_n^{**}}(\theta^{**}, 1)}{Z_{n,\theta_n^{**}}(\theta^{**}, t_o)}$ est borné en probabilité

par conséquence du théorème 1, le comportement de $\tilde{Z}_{n,\theta_n^*}(\theta^{**} + \sqrt{n}\, d_n, t_o - \frac{k}{n})$

est réglé par le lemme (3.6) (1).

Sur $B_{2,n}$: $\dfrac{Z_{n,\theta_n^{**}}(\theta^{**}, 1)}{Z_{n,\theta_n^{**}}(\theta^{**}, t_o)}$ est borné en probabilité, le comportement

de $\tilde{Z}^f_{n,\theta_n^*}(\theta^{**} + \sqrt{n}\, d_n, t_o - \frac{k}{n})$ est réglé par le lemme (3.6) (2) ,

celui de $Z_{n,\theta_n^*}(\theta^*, \frac{k}{n})$ par la proposition 1.1. .

Enfin sur $B_{3,n}$, les comportements de $\tilde{Z}_{n,\theta_n^*}(\theta^{**} + \sqrt{n}\, d_n, t_o - \frac{k}{n})$ de $Z_{n,\theta_n^*}(\theta^*, \frac{k}{n})$

sont réglés par la proposition 1.1., la compensation étant apportée par

$\dfrac{Z_{n,\theta_n^{**}}(\theta^{**}, 1)}{Z_{n,\theta_n^{**}}(\theta^{**}, t_o)}$ à l'aide du lemme 4.5. .

Nous pouvons maintenant déduire la loi asymptotique des estimateurs du maximum de vraisemblance :

Théorème (3.3)

Sous les conditions du théorème (3.2), les estimateurs du maximum de vraisemblance des paramètres $\bar{\theta}_o$, d_n, t_o sont asymptotiquement indépendants et de lois respectives :

$$\sqrt{n}(\hat{\bar{\theta}}_n - \bar{\theta}_o) \xrightarrow[n\to\infty]{} N(0, \mathcal{J}^{-1}_{[0,1]}(\sigma_o))$$

$$\sqrt{n}(\hat{d}_n - d_n) \xrightarrow[n\to\infty]{} N(0, A_{t_o}(\sigma_o))$$

avec $A_{t_o}(\sigma_o) = (\mathcal{J}_{[o,t_o]}(\sigma_o) - \mathcal{J}_{[o,t_o]}(\sigma_o)\mathcal{J}^{-1}_{[0,1]}(\sigma_o)\mathcal{J}_{[0,t_o]}(\sigma_o))^{-1/2}\mathcal{J}^{-1}_{[0,1]}(\sigma_o) \times$

$$(\mathcal{J}_{[0,t_o]}(\sigma_o) - \mathcal{J}_{[0,t_o]}(\sigma_o)\mathcal{J}^{-1}_{[0,1]}(\sigma_o)\mathcal{J}_{[0,t_o]}(\sigma_o))^{-1/2}$$

et $(d_n' \mathcal{F}(t_o)' \mathcal{J}(\sigma_o) \mathcal{F}(t_o) d_n)(\hat{k} - nt_o) \xrightarrow[n\to\infty]{} \Gamma$ où Γ admet la densité sur $\mathbb{R}$: $\gamma(t) = \{\frac{3}{2} e^{|t|} \Phi(\frac{3}{\sqrt{2}}\sqrt{|t|}) - \frac{1}{2}\Phi(\frac{1}{2}\sqrt{|t|}\}$

avec $\Phi(t) = \int_t^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$.

Démonstration du théorème

Désignons par $(\hat{T}_n, \hat{\bar{\theta}}_n, \hat{U}_n)$ la suite des estimateurs du maximum de vraisemblance calculés sur le processus Λ^2_n .

Soit B un pavé fermé borné quelconque de $\mathbb{R}^{2p+1}$ on a :

$$P((\hat{T}_n, \hat{\bar{\theta}}_n, \hat{U}_n) \in B) =$$

$$P(\sup_{(T,\bar{\theta},U)\in B} \Lambda^2_n(T,\bar{\theta},U) - \sup_{(T,\bar{\theta},U)\in \mathbb{R}^{2p+1}\setminus B} \Lambda^2_n(T,\bar{\theta},U) > 0)$$

de la continuité dans $\mathcal{C}_o$ de la fonctionnelle $h \longrightarrow \sup_B h - \sup_{\mathbb{R}^{2p+1}\setminus B} h$,

on déduit par le théorème (3.2) la convergence de l'expression ci-dessus

vers $P(\sup_{(T,\bar{\theta},U)\in B} \Lambda^2(T,\bar{\theta},U) - \sup_{(T,\bar{\theta},U)\in \mathbb{R}^{2p+1}\setminus B} \Lambda^2(T,\bar{\theta},U) > 0)$

$= P((\hat{T},\hat{\bar{\theta}},\hat{U}) \in B)$ pourvu que l'instant d'atteinte du suprêmum $(\hat{T},\hat{\bar{\theta}},\hat{U})$ soit presque sûrement unique .

La seule partie non triviale à ce stade concerne $\hat{T}$:

En utilisant les résultats sur les instants de suprêmum d'un Brownien (voit Shepp [17]), nous pouvons écrire la densité du triplet $(Y,X,\hat{T})$ où $Y = \sup_{T\in[0,\tau]} W(T) - \frac{|T|}{2}$, $X = W(\tau) - \frac{\tau}{2}$ et $\hat{T}$ l'instant où le suprémum est atteint sur $[0,\tau]$:

$$\frac{1}{\pi} \frac{y(y-x)}{\hat{T}^{3/2}(\tau-\hat{T})^{3/2}} e^{-\frac{y^2}{2\hat{T}} - \frac{(y-x)^2}{2(\tau-\hat{T})}} \cdot e^{-\frac{x}{2}-\frac{\tau}{8}} 1_{[0,\tau]}(\hat{T})\, 1_{[0,\infty[}(y)\, 1_{[0,\infty[}(y-x)$$

la densité du couple $(y,\hat{T})$ dans les mêmes conditions vaut :

$$\frac{1}{\pi} \frac{y}{\hat{T}^{3/2}} e^{-\frac{1}{2}(\frac{y}{\sqrt{\hat{T}}} + \frac{\sqrt{\hat{T}}}{2})^2} 1_{[0,\tau]}(\hat{T})\, 1_{[0,\infty[}(y) \left[\frac{e^{-\frac{(\tau-\hat{T})}{8}}}{\sqrt{\tau}-\hat{T}} + \frac{\sqrt{2\pi}}{2} \Phi(-\frac{\sqrt{\tau-\hat{T}}}{2})\right]$$

lorsque τ tend vers l'infini, cette loi converge vers la densité :

$$\frac{1}{\sqrt{2\pi}} \frac{y}{\hat{T}^{3/2}} e^{-\frac{1}{2}(\frac{y}{\sqrt{\hat{T}}} + \frac{\sqrt{\hat{T}}}{2})^2} 1_{[0,\infty[}(\hat{T})\, 1_{[0,\infty[}(y)$$

Pour établir la loi de l'instant de rupture, il suffit de considérer $(y_1,\hat{T}_1)$ et $(y_2,\hat{T}_2)$ indépendants et admettant la densité ci-dessus ($\hat{T}_1 > 0$ et $\hat{T}_2 < 0$) et de calculer la densité de $\hat{T} = \hat{T}_1\, 1_{y_1\geq y_2} + \hat{T}_2 1_{y_1<y_2}$ qui pour $\hat{T}$ positif vaut :

$$\gamma(\hat{T}) = \gamma(-\hat{T}) = \int_0^{\infty} \frac{y_1}{\sqrt{2\pi}} \; e^{-\frac{1}{2}\left(\frac{y_1}{\sqrt{|\hat{T}|}} + \frac{\sqrt{|\hat{T}|}}{2}\right)} \int_0^{y_1} dy_2 \; e^{-y_2} \; dy_1 \; .$$

Par un calcul simple, on obtient le résultat annoncé.

IV.- DEMONSTRATION DU THEOREME 1

Celle-ci comportera trois parties :

1.- Convergence uniforme de Z_n^f sur les compacts lorsque f est une fonction en escalier : lemme (4.1).

2.- Approximation uniforme sur les compacts de Z_n^f par Z_n^e où e est une fonction matricielle en escalier : lemme (4.2).

3.- Convergence de Z_n^f dans $\mathscr{C}_\Phi$.

Lemme (4.1)

Sous les conditions du théorème 1.1., pour toute fonction matricielle e en escalier sur $[0,1]$, la suite de processus $Z_n^e(.,.)$ converge dans $\mathscr{C}_\Phi(\mathbb{R}^p \times [0,1])$ sous $H_{o,n}(\sigma_o)$ uniformément sur les compacts en σ_o vers le processus $Z^e(.,.)$ défini dans l'énoncé du théorème 1.

Démonstration : La fonction e étant en escalier, sur chaque intervalle de $[0,1]$ où e est constante toutes les conditions du théorème 2 de [21] sont remplies et le lemme ci-dessus en est une conséquence immédiate.

Lemme (4.2)

Sous les hypothèses A à F, pour tout ε positif et pour tout compact K de $\mathbb{R}^p$, il existe δ positif tel que pour toute fonction en escalier e satisfaisant : $\sup_{t\in[0,1]} \|f(t) - e(t)\| < \delta$, il existe $n_o(\varepsilon,K,e)$ tel que pour tout n supérieur à n_o :

$$P\{\sup_t \sup_{\theta\in K} |Z_n^f(\theta,t) - Z_n^e(\theta,t)| > \varepsilon\} \leqslant \varepsilon$$

Démonstration

Remarquons d'abord que les fonctions f et e étant bornées, l'hypothèse B étant satisfaite, une démonstration strictement analogue à celle du lemme (2.8) de [21] implique que les deux suites de processus $Z_n^f(.;1)$ et $Z_n^e(.;1)$ sont tendues dans $\mathscr{C}(\mathbb{R}^p)$ muni de la topologie de la convergence uniforme sur les compacts. Avec l'hypothèse A, on en déduit la convergence étroite dans $\mathscr{C}(\mathbb{R}^p)$ de la suite des couples $(Z_n^f(.;1), Z_n^e(.;1))$.

Maintenant étudions la différence des processus $Z_n^f(.;.) - Z_n^e(.;.)$:

$$P\{\sup_{t\in[0,1]} \sup_{\theta\in K} |Z_n^f(\theta,t) - Z_n^e(\theta,t)| > \varepsilon\}$$

$$< P\{\sup_{t\in[0,1]} \sup_{\theta\in K} \left|\frac{Z_n^f(\theta,t)}{Z_n^e(\theta,t)} - 1\right| > \frac{\varepsilon}{A}\} + P\{\sup_{t\in[0,1]} \sup_{\theta\in K} Z_n^e(\theta,t) > A\}$$

D'après le lemme (4.1), pour tout A :

$P\{\sup_{t\in[0,1]} \sup_{\theta\in K} Z_n^e(\theta,t) > A\}$ converge vers :

$$P\{\sup_{t\in[0,1]} \sup_{\theta\in K} Z^e(\theta,t) > A\}$$

$$\leqslant P\{\sup_{t\in[0,1]} \left\| \int_0^t \mathscr{E}(s)\, \mathfrak{I}^{1/2}(\sigma_0)\, dW(s) \right\| > \frac{A}{\|K\|}\}$$

ce dernier terme étant inférieur à $\frac{\varepsilon}{4}$ pour $A > A_0(\varepsilon, \|e\|)$.

Le premier terme se majore à l'aide du lemme (1.6) de [21] :

$$P\{\sup_{k=1,\ldots,n} \sup_{\theta\in K} \left|\log \frac{Z_n^f(\theta,\frac{k}{n})}{Z_n^e(\theta,\frac{k}{n})}\right| \geqslant \frac{\varepsilon}{A_0}\}$$

$$\leqslant P\{\sup_{\theta\in K} \left|\log \frac{Z_n^f(\theta,1)}{Z_n^e(\theta,1)}\right| > \frac{\varepsilon}{2A_0}\} \times \left[1 - \sup_{k=1\ldots n} P\{\sup_{\theta\in K} \left|\log \frac{\overleftarrow{Z^f}(\theta,\frac{k}{n})}{\overleftarrow{Z^e}(\theta,\frac{k}{n})}\right| > \frac{\varepsilon}{2A_0}\}\right]^{-1}$$

où $\overleftarrow{f}(t) = f(1-t)$, $\overleftarrow{e}(t) = e(1-t)$.

Les processus limites $Z^f(.;1)$ et $Z^e(.;1)$ ne chargeant pas 0 et l'infini ,

$$P\{\sup_{\theta\in K} \left|\log \frac{Z_n^f(\theta,1)}{Z_n^e(\theta,1)}\right| > \frac{\varepsilon}{2A_o}\} \text{ converge vers } P\{\sup_{\theta\in K} \left|\log \frac{Z^f(\theta,1)}{Z^e(\theta,1)}\right| > \frac{\varepsilon}{2A_o}\}$$

quantité qui est majoré par $\frac{\varepsilon}{4}$ si

$$\left\|\int_o^1 \left[\mathcal{F}(s) - \mathcal{E}(s)\right]' \mathfrak{I}(\sigma_o)\left[\mathcal{F}(s) - \mathcal{E}(s)\right]ds\right\| \leqslant \delta_1(\varepsilon) \text{ ce qui est}$$

réalisé en particulier dès que $\sup_t \|f(t) - e(t)\| \leqslant \delta_2(\varepsilon)$.

Il reste à montrer que :

$$\sup_{k=1,\ldots,n} P\{\sup_{\theta\in K} \left|\log \frac{Z_n^{\overleftarrow{f}}(\theta,\frac{k}{n})}{Z_n^{\overleftarrow{e}}(\theta,\frac{k}{n})}\right| > \frac{\varepsilon}{2A_o}\} \leqslant \frac{1}{2}$$

$$\log \frac{Z_n^{\overleftarrow{f}}(\theta;\frac{k}{n})}{Z_n^{\overleftarrow{e}}(\theta;\frac{k}{n})} = - \sum_{i=1}^{k} \frac{\eta'}{\sqrt{n}} \left[\overleftarrow{f}(\tfrac{i}{n}) - \overleftarrow{e}(\tfrac{i}{n})\right]' \dot{L}_x(X_i ; \sigma_o + \frac{\sigma}{\sqrt{n}})$$

$$+ \sum_{i=1}^{k} \frac{\eta'}{\sqrt{n}} \left[\overleftarrow{f}(\tfrac{i}{n}) - \overleftarrow{e}(\tfrac{i}{n})\right]' \ddot{L}_{xx}(X_i - \tau_1 \overleftarrow{f}(\tfrac{i}{n})\eta ; \sigma_o + \frac{\sigma}{\sqrt{n}}) \overleftarrow{f}(\tfrac{i}{n})\frac{\eta}{\sqrt{n}}$$

$$- \frac{1}{2} \sum_{i=1}^{k} \frac{\eta'}{\sqrt{n}} \left[\overleftarrow{f}(\tfrac{i}{n}) - \overleftarrow{e}(\tfrac{i}{n})\right]' \ddot{L}_{xx}(X_i - (\tau_2 \overleftarrow{f}(\tfrac{i}{n}) + (1-\tau_2)\overleftarrow{e}(\tfrac{i}{n}))\frac{\eta}{\sqrt{n}} ; \sigma_o + \frac{\sigma}{\sqrt{n}})$$

$$\left[\overleftarrow{f}(\tfrac{i}{n}) - \overleftarrow{e}(\tfrac{i}{n})\right] \frac{\eta}{\sqrt{n}}$$

avec $0 < \tau_1 < 1$ et $0 < \tau_2 < 1$ et $\dot{L}_x(., \sigma_o + \frac{\sigma}{\sqrt{n}})$, $\ddot{L}_{xx}(., \sigma_o + \frac{\sigma}{\sqrt{n}})$

désignant les gradients et matrices Jacobienne et Hessienne en x de $L(., \sigma_o + \frac{\sigma}{\sqrt{n}})$.

La loi des grands nombres implique que le second terme est majoré par :

$$\|K\|^2 \, \|f\| \, \|f-e\| E \sup_{\|\mu\|+\|\sigma\|<\delta_o} \|\overset{\bullet\bullet}{L}_{xx}(X-\mu \, ; \, \sigma_o+\sigma)\| \quad \text{si on note}$$

$$\|K\| = \sup_{\theta \in K} \|\theta\|$$

et le troisième terme par : $\frac{1}{2}\|K\|^2 \, \|f-e\|^2 \, E \sup_{\|\mu\|+\|\sigma\|\leq\delta_o} \|\overset{\bullet\bullet}{L}_{xx}(X-\mu \, ; \, \sigma_o+\sigma)\|$

avec une probabilité $\geq 1 - \frac{\varepsilon}{4}$ dès que $k \geq k_o(\delta_o,\varepsilon)$ et $n \geq \left(\frac{\|f\| \cdot \|K\|}{\delta_o}\right)^2$

chacun de ces termes sera donc rendu $< \frac{\varepsilon}{6A_o}$ avec une probabilité supérieure à $1 - \frac{\varepsilon}{4}$ si $k \geq k_o(\delta_o,\varepsilon)$ et $\|f-e\| < \delta$ suffisamment petit.

Le premier terme est majoré par (les notation allant de soi)

$$\|K\| \cdot \|f-e\| \cdot \| \frac{1}{\sqrt{n}} \sum_{i=1}^{k} \dot{L}_x(X_i, \, \sigma_o + \frac{\sigma}{\sqrt{n}}) \| \qquad (4.0)$$

Or, on a : $\| \frac{1}{\sqrt{n}} \sum_{i=1}^{k} \dot{L}_x(X_i, \, \sigma_o + \frac{\sigma}{\sqrt{n}}) \|$

$$\leq \frac{1}{\sqrt{n}} \| \sum_{i=1}^{k} \dot{L}_x(x_i,\sigma_o) \| + \frac{1}{n} \| \sum_{i=1}^{k} \tilde{L}_{x\sigma}(x_i, \, \sigma_o + \frac{\tilde{\sigma}}{\sqrt{n}}) \|$$

Une nouvelle application de la loi des grands nombres et l'utilisation de l'inégalité maximale pour le 1er terme de cette dernière somme montre que (4.0) peut être rendu inférieur $\frac{\sigma}{6A_o}$ avec une probabilité supérieure $1 - \frac{\varepsilon}{4}$ si $k \geq k_o(\delta_o,\varepsilon)$. On obtient alors :

$$P\{\sup_{\theta \in K} \left| \log \frac{\overleftarrow{Z}_n^f(\theta,\frac{k}{n})}{\overleftarrow{Z}_n^e(\theta,\frac{k}{n})} \right| > \frac{\varepsilon}{2A_o}\} < \frac{\varepsilon}{2} \quad \text{si} \quad k \geq k_o(\delta_o,\varepsilon), \; n > \left(\frac{\|f\| \cdot \|K\|}{\delta_o}\right)^2$$

et $\|f-e\| < \delta$.

Pour les k_o premiers termes, on a :

$$\sup_{k=1,\dots,k_o} P\{\sup_{\theta\in K} \sum_{i=1}^{k} \log \frac{g(X_i - \overleftarrow{f}(\frac{i}{n})\frac{\eta}{\sqrt{n}}; \sigma_o + \frac{\sigma}{\sqrt{n}})}{g(X_i - \overleftarrow{e}(\frac{i}{n})\frac{\eta}{\sqrt{n}}, \sigma_o + \frac{\sigma}{\sqrt{n}})} > \frac{\varepsilon}{2A_o}\}$$

$$< \sum_{i=1}^{k_o} P\{\sup_{\theta\in K} \left| \log \frac{g(X - \overleftarrow{f}(\frac{i}{n})\frac{\eta}{\sqrt{n}}; \sigma_o + \frac{\sigma}{\sqrt{n}})}{g(X - \overleftarrow{e}(\frac{i}{n})\frac{\eta}{\sqrt{n}}; \sigma_o + \frac{\sigma}{\sqrt{n}})} \right| > \frac{\varepsilon}{2A_o . k_o}\}$$

$$< k_o . P\{\sup_{\|(\mu,\sigma)\|<\delta_o} \|\dot{L}_x(X-\mu;\sigma_o+\sigma)\| \cdot \|f-e\| \cdot \frac{\|K\|}{\sqrt{n}} > \frac{\varepsilon}{2A_o\, k_o}\}$$

arbitrairement petit si $\|f-e\|$ est choisi en conséquence.

Soit ϕ la fonction matricielle possédant les propriétés énoncées au paragraphe 1 et vérifiant la condition (1.0), la convergence de Z_n^f dans $\mathscr{C}_\Phi$ revient à la majoration uniforme :

$$\forall \varepsilon > 0,\ \exists M_o(\varepsilon) \text{ tel que } \forall M \geq M_o(\varepsilon),\ \forall n$$

$$P_{\theta_n}\{\sup_{t\in]0,1]} \sup_{\theta,\ \theta\phi(t)\theta > M} Z_n^f(\theta,t) \geq \varepsilon\} \leq \varepsilon$$

Quitte à changer M en $\frac{M}{C}$, la condition (1.0) permet de montrer la condition plus forte où on agrandit le domaine où on prend le suprémum

$$P_{\theta_n}\{\sup_{t\in]0,1]} \sup_{\theta,\ \theta' \frac{\overline{F}(t)}{t}\theta > \frac{M}{t^{1+\alpha}}} Z_n^f(\theta,t) \geq \varepsilon\} \leq \varepsilon \qquad (4.1)$$

Cette dernière inégalité repose fondamentalement sur une succession de lemmes qui ont été montrés en détail dans [21] pour une suite de processus Z_n en l'absence de régression et que l'on peut récrire point par point:

Lemme (4.3)

Sous les hypothèses A, C, F , on a :

$\forall N > 0$, $\forall R > 0$, $\exists B(N,R) > 0$ et $\exists n_o(N,R)$ tels que :

$$\forall \sigma_o,\ \|\sigma_o\| < R,\ \forall \theta \in \mathbb{R}^p,\ \forall t \in]0,1],\ \forall n,\ [nt] \geqslant n_o,\ E_{\sigma_o} |Z_n^f(\theta;t)|^{\frac{q}{2}} < \frac{\Gamma \cdot B(N,R)}{|\theta' \bar{F}(t)\theta|^{\frac{\gamma N}{2}}}$$

Démonstration

Une majoration identique a été démontrée dans le lemme (2.1) de [21] en l'absence de régression avec en supplément ici l'uniformité en σ_o, $\|\sigma_o\| < R$ qui découle des hypothèses A ii) et C :

$$E_{\sigma_o} [Z_n((\mu,\sigma);1)]^{\frac{q}{2}} < \frac{B(N,R)}{\|(\mu,\sigma)\|^N} \wedge 1$$

on en déduit pour le processus avec régression :

$$E_{\sigma_o} [Z_n^{f(.)}(\theta,\tfrac{k}{n})]^{\frac{q}{2}} = E_{\sigma_o} [Z^{f(\frac{k}{n}\bullet)}(\theta\sqrt{\tfrac{k}{n}};1)]^{\frac{q}{2}}$$

$$< \Big(\prod_{i=1}^{k} \frac{B(N,R)}{\|\sqrt{\tfrac{k}{n}}\, \mathcal{F}(\tfrac{k}{n}\bullet\tfrac{i}{k})\theta\|^N}\Big)^{\frac{1}{k}}$$

La condition F fournit un bon encadrement de $\mathcal{F}(s)\theta$:

$$\eta' \frac{F(t)}{t} \eta \leqslant \sup_{s \leqslant t} \|f(s)\eta\|^2 \leqslant \frac{\Gamma^2}{\gamma} \cdot \eta' \frac{F(t)}{t} \eta$$

ce qui implique :

$$\theta' \bar{F}(t)\theta \leqslant t \sup_{s \leqslant t} \|\mathcal{F}(s)\theta\|^2 \leqslant \frac{\Gamma^2}{\gamma}\ \theta' \bar{F}(t)\theta \qquad (4.2)$$

on achève ainsi la démonstration du lemme (4.3) :

$$E_{\sigma_o} |Z_n^f(\theta,\tfrac{k}{n})|^{\frac{q}{2}} < \frac{\Gamma \cdot B(N,R)}{\sup_{i=1,\ldots,k} \| \sqrt{\tfrac{k}{n}}\, \mathcal{F}(\tfrac{i}{n})\theta \|^{\gamma N}} < \frac{\Gamma \cdot B(N,R)}{\left[\theta' \bar{F}(\tfrac{k}{n})\theta\right]^{\frac{\gamma N}{2}}}$$

En reprenant les notations u, $\bar{u}$, $\ell(u)$, $\pi_u(\theta)$, θ^u, $[\theta_1,\theta_2]$, ... qui ont été introduites pour démontrer les lemmes (2.7) et (2.8) de [21], on peut énoncer le :

Lemme (4.4)

Sous les hypothèses A,B,C et F, $\forall N > 0$, $\forall R > 0$, $\exists B(N,R)$, $\exists n_o(N)$ tels que : $\forall u \subseteq \{1,\ldots,p\}$, $\forall \sigma_o, \|\sigma_o\| \leqslant R$, $\forall \theta_1$, $\forall \theta_2$, $\forall \theta_3 \in \mathbb{R}^p$, $\forall t \in]0,1]$, $\forall n$, $[nt] \geqslant n_o$

$$E_{\sigma_o}\left[Z_n^f(\theta_3^{\bar{u}}, \pi_u[\theta_1,\theta_2] ; t)\right]^{q'} \leqslant B(N,R) \frac{\sup_{\theta \in \mathcal{U}(\theta_1,\theta_2,\theta_3)} \left[\theta' \bar{F}(t)\theta\right]^{a'}}{\inf_{\theta \in \mathcal{U}(\theta_1,\theta_2,\theta_3)} \left[\theta' \bar{F}(t)\theta\right]^{N}} \cdot \lambda_u(\pi_u|\theta_1\sqrt{t},\ \theta_2\sqrt{t}|)^{\alpha'_u}$$

avec $q' \in \left]\frac{q}{2}, q\right[$ choisi de sorte que $\alpha'_u > 1$ (c'est-à-dire $m.q' > p$)

$$\mathcal{U}(\theta_1,\theta_2,\theta_3) = \{\theta \in \mathbb{R}^p,\ \pi_u(\theta) \in \pi_u[\theta_1,\theta_2] \text{ et } \pi_{\bar{u}}(\theta) = \theta_3^{\bar{u}}\}$$

Démonstration

L'extension stricte du lemme (2.7) de [21] fournit :

$$E_{\sigma_o} |Z_n^f(\theta_3^{\bar{u}}, \pi_u[\theta_1,\theta_2];1)|^q \leqslant B(R) \left[1 \vee \sup_{s \in [0,1]} \sup_{\theta \in \mathcal{U}(\theta_1,\theta_2,\theta_3)} \left\|\mathcal{F}(s)\theta\right\|\right]^{a'} \lambda_u(\pi_u[\theta_1,\theta_2])^{\alpha_u}$$

et l'inégalité de Hôlder utilisée avec le résultat du lemme (4.3) donne le résultat annoncé.

L'intérêt majeur de ce dernier lemme est qu'il fournit un contrôle à l'infini du processus $Z_n^f(.,t)$ à t fixé :

Lemme (4.5)

$\forall N > 0$, $\forall R > 0$, $\exists B'(N,R)$, $\exists B''(N,R)$ et $\exists n_o(N)$ tels que $\forall A > 0$, $\forall \sigma_o$, $\|\sigma_o\| \leqslant R$, $\forall \theta_1$, $\forall \theta_2 \in \mathbb{R}^p$, $\|\theta_1 - \theta_2\| \leqslant 1$, $\forall t \in]0,1]$, $\forall n$, $[nt] \geqslant n_o$,

$$P_{\sigma_o}\left\{ \sup_{\substack{\theta \\ \sqrt{t}\,\theta \in [\theta_1,\theta_2]}} Z_n^f(\theta,t) \geqslant A \right\} < \frac{B'}{A^{q'}} \cdot \frac{\sup\limits_{\theta \in [\theta_1,\theta_2]} \left[\theta' \bar{F}(t) \theta\right]^{a'}}{\inf\limits_{\theta \in [\theta_1,\theta_2]} \left[\theta' \bar{F}(t) \theta\right]^{N}} \qquad (4.3)$$

et $\forall \ell \geqslant 4 \|\bar{F}(t)\|$,

$$P_{\sigma_o}\left\{ \sup_{\theta,\ \theta \bar{F}(t) \theta > \ell} Z_n^f(\theta,t) \geqslant A \right\} < \frac{B''}{A^{q'}} \left| \det \frac{\bar{F}(t)}{t} \right|^{-\frac{1}{2}} \ell^{a' + \frac{p}{2} - N} \qquad (4.4)$$

Démonstration

La première assertion (4.3) est une conséquence du lemme de CSENSOV (voir le lemme (1.3) dans [21]) et du lemme (4.4).

La seconde assertion (4.4) s'obtient à partir de la précédente en sommant sur les cubes élémentaires $[\theta_I, \theta_{I+1}]$ de côtés $\frac{1}{\sqrt{t}}$:

$$P_{\sigma_o}\left\{ \sup_{\theta,\ \theta \bar{F}(t) \theta > \ell} Z_n^f(\theta,t) \geqslant A \right\} < \frac{B'}{A^{q'}} \cdot \sum_{I,\ \theta_I' \bar{F}(t) \theta_I > \ell} \frac{\sup\limits_{\theta \in [\theta_I,\theta_{I+1}]} \left[\theta' \bar{F}(t) \theta\right]^{a'}}{\inf\limits_{\theta \in [\theta_I,\theta_{I+1}]} \left[\theta' \bar{F}(t) \theta\right]^{N}}$$

$$< \frac{B''}{A^{q'}} \sum_{j=\ell}^{\infty} j^{a'-N} \{\text{Card } I,\ j \leqslant \theta_J' \bar{F}(t) \theta_J < j+1\}$$

$$< \frac{B''}{A^{q'}} \sum_{j=\ell}^{\infty} j^{a'-N+\frac{p}{2}-1} \left[\det \frac{\bar{F}(t)}{t}\right]^{-\frac{1}{2}} .$$

Ayant le contrôle du processus $Z_n^f(.,t)$ à t fixé, il reste à obtenir l'uniformité en t, c'est-à-dire la convergence de Z_n^f dans $\mathscr{C}_\Phi$ ou mieux : l'inégalité (4.1) où on découpe le domaine où est pris le suprémum en deux parties :

$$C_1(M) = \{(k,\theta) \in \{1,\dots,n\} \times \mathbb{R}^p \ , \ (\tfrac{k}{n})^{-\alpha} M \leqslant \theta' \overline{F}(\tfrac{k}{n})\theta \leqslant k\delta^2\}$$

$$C_2(M) = \{(k,\theta) \in \{1,\dots,n\} \times \mathbb{R}^p \ , \ \theta' \overline{F}(\tfrac{k}{n})\theta \geqslant (\tfrac{k}{n})^{-\alpha} M \vee k\delta^2\}$$

et on aura ainsi à vérifier les deux inégalités :

$\forall \varepsilon > 0$, $\forall \delta > 0$, $\exists M_o(\varepsilon,\delta)$ tel que $\forall M \geqslant M_o$, $\forall n$

$$P\{\sup_{C_2(M)} Z_n^f(\theta,\tfrac{k}{n}) \geqslant \varepsilon\} \leqslant \varepsilon \qquad (*)$$

$\forall \varepsilon > 0$, $\exists \delta_o(\varepsilon)$ et $\exists M_o(\varepsilon)$ tels que $\forall M \geqslant M_o$, $\forall \delta \leqslant \delta_o$, $\forall n$

$$P\{\sup_{C_1(M)} Z_n^f(\theta,\tfrac{k}{n}) \geqslant \varepsilon\} \leqslant \varepsilon \qquad (**)$$

Montrons d'abord l'inégalité (*). Ayant fixé M arbitrairement grand, faisons la majoration pour $n \geqslant \dfrac{M}{C\delta^2}$ et notons $m = \left[(\dfrac{M}{\delta^2} n^\alpha)^{\frac{1}{1+\alpha}}\right]$.

On décompose le 1er membre de (*) en 3 parties : $\pi_1 + \pi_2 + \pi_3$ avec :

$$\pi_1 = P\Big(\sup_{C_2(M) \cap \{k \leqslant n_o\}} Z_n^f(\theta,\tfrac{k}{n}) \geqslant \varepsilon\Big) \ ,$$

$$\pi_2 = \sum_{k=n_o}^{m} P\Big(\sup_{C(\frac{k}{n}, \frac{M}{(\frac{k}{n})^\alpha})} Z_n^f(\theta,\tfrac{k}{n}) \geqslant \varepsilon\Big) \ ,$$

$$\pi_3 = \sum_{k=m+1}^{n} P\Big(\sup_{C(\frac{k}{n}, k\delta^2)} Z_n^f(\theta,\tfrac{k}{n}) \geqslant \varepsilon\Big)$$

où $C(t,\ell) = \{\theta \in \mathbb{R}^p \ , \ \theta' \overline{F}(t)\theta > \ell\}$.

On obtient par l'inégalité (4.4) :

$$\pi_2 \leq \sum_{k=n_o}^{m} \frac{B'}{\sqrt{\det(\frac{\overline{F}(\frac{k}{n})}{\frac{k}{n}})}} \left[\frac{(\frac{k}{n})^{\alpha}}{M}\right]^{N-a'-\frac{p}{2}}$$

et en prenant l'équivalent de $\overline{F}$ au voisinage de 0 , on obtient l'existence de $s \geq 0$ tel que :

$$\pi_2 \leq \frac{B}{M^{N-a'-\frac{p}{2}}} \sum_{k=n_o}^{m} \frac{1}{(\frac{k}{n})^{s-\alpha(N-a'-\frac{p}{2})}}$$

$$\leq \frac{B}{M^{N-a'-\frac{p}{2}}} \ m. \ (\frac{m}{n})^{\alpha(Na'-\frac{p}{2})-s} \qquad \text{dès que } \alpha(N-a'-\tfrac{p}{2}) > s$$

$$\leq \frac{B}{M^{N-a'-\frac{p}{2}}} (\frac{M}{\delta^2})^{\frac{1}{1+\alpha}} + \frac{N-a'-\frac{p}{2}}{1+\alpha} \ n^{\frac{\alpha[1-N+a'+\frac{p}{2}]}{1+\alpha}}$$

$$\leq \varepsilon \quad \text{pour} \quad N-a'-\tfrac{p}{2} > 1 \quad \text{et} \quad n > n_o(\varepsilon).$$

De même,

$$\pi_3 < B' \sum_{k=m+1}^{n} \frac{1}{\sqrt{\det \frac{\overline{F}(\frac{k}{n})}{\frac{k}{n}}}} (\frac{1}{k\delta^2})^N$$

$$\pi_3 \leq B' \frac{n^s}{\delta^{2N} \ m^{N+s-1}}$$

$$\leq \frac{B'}{\delta^{2N}} (\frac{M}{\delta^2})^{-\frac{N+s-1}{1+\alpha}} \ n^{s-\frac{\alpha(N+s-1)}{1+\alpha}}$$

$$\leq \varepsilon \quad \text{dès que} \quad n > n_o(\alpha) \quad \text{et} \ s < \frac{\alpha(N+s-1)}{1+\alpha}$$

On a en utilisant l'inégalité (4.2)

$$\pi_1 < \sum_{k=1}^{n_o} P\{ \sup_{C(\frac{k}{n}, \frac{M}{(\frac{k}{n})^\alpha} \vee k\delta^2)} \prod_{i=1}^{k} \frac{g(X_i - f(\frac{i}{n})\frac{\eta}{\sqrt{n}}, \sigma_o + \frac{\sigma}{\sqrt{n}})}{g(X_i, \sigma_o)} \geqslant \varepsilon^q \}$$

$$\leqslant n_o P(\sup_{\|(\mu\ \sigma)\|^2 > M} \frac{g(X_i - \mu, \sigma_o + \sigma)}{g(X_i, \sigma_o)} \geqslant \varepsilon^{\frac{q}{n_o}})$$

ce dernier terme peut être rendu inférieur à ε pour M sufisamment grand grâce à l'hypothèse D, ce qui achève la démonstration de (*)

Montrons pour finir, l'inégalité (**) à l'aide du lemme suivant :

Soir $k_o(n)$ un entier de $[0,n]$, A_n uns constante positive, h une fonction continue de $]0,1]$ dans $\mathbb{R}_*^+$ tel que $H(t) = \frac{h(t)}{t}$ soit strictement croissante et B_h l'ensemble :

$$B_h = \{\theta \ / \ \frac{M}{h(\frac{k}{n})} \leqslant \frac{\theta' \overline{F}(\frac{k}{n})\theta}{\frac{k}{n}} \leqslant n\delta^2\}$$

Lemme (4.6)

Sous la condition D, on a : Pour tout ε positif, il existe n_o et des constante positives C_1, C_2 tels que $n > n_o$ implique :

$$P_{(0,\sigma_o^2)} (\sup_{k_o'(n) \leqslant k \leqslant n} \sup_{\theta \in B_h} Z_n^f(\theta, \frac{k}{n}) > A_n) < \varepsilon +$$

$$C_1 \frac{|\log \frac{k_o'(n)}{n}|}{\log A_n^q} 1_{[A_n > 1]} + \frac{C_2}{M} \int_{\frac{K(n)}{n}}^1 \frac{h(t)}{t^2} dt \quad \text{où}$$

$$\frac{K(n)}{n} = H^{-1}\left(\frac{\frac{1-\varepsilon}{2}\,\underline{\lambda}(\sigma_o)M}{\log A_n^q} \vee 0\right), \quad k_o'(n) = \max\{n_o, k_o(n)\}$$

et $\underline{\lambda}(\sigma_o) = \dfrac{1}{\| \mathcal{J}(\sigma_o)^{-1} \|}$.

Démonstration

$$P_{(0,\sigma_o)}\left\{ \sup_{k_o(n)\leq k\leq n} \sup_{\theta\in B_h} Z_n^f(\theta,\tfrac{k}{n}) > A_n \right\} \leq$$

$$P\Big\{ \sup_{k_o(n)\leq k\leq n} \sup_{\theta\in B_h} \frac{\theta'}{\sqrt{n}} \sum_{i=1}^{k} \mathscr{F}(\tfrac{i}{n})' \dot{L}(X_i,\sigma_o) + \frac{1}{2n}\,\theta' \sum_{i=1}^{k} \mathscr{F}(\tfrac{i}{n})' \cdot \ddot{L}(X_i - f(\tfrac{i}{n})\frac{\hat{\eta}_k}{\sqrt{n}}, \sigma_o + \frac{\hat{\sigma}_k}{\sqrt{n}})\, \mathscr{F}(\tfrac{i}{n})\theta > \log A_n^q \Big\}$$

avec $\hat{\theta}_k$ appartenant à $[0,\theta]$ et $\dot{L}(X_i,\sigma_o)$ désigne le gradient de $\log g(X_i - f(\frac{i}{n})\eta,\sigma)$ au point $(0,\sigma_o)$.

Le controle de la partie quadratique en θ va être assuré grâce à une loi des grands nombres et il est utile de définir maintenant $\overline{\lambda}$ et $\underline{\lambda}$ les plus grandes et plus petites valeurs propres de la matrice carrée symétrique $\ddot{L}$ et pour tout δ positif (petit), le théorème de section mesurable (qu nous avons utilisé dans [21]) permet d'exhiber des versions mesurables $T_s(X,\delta) = (\mu_s(X,\delta), \sigma_s(X,\delta))$ et $T_i(X,\delta) = (\mu_i(X,\delta), \sigma_i(X,\delta))$ à valeurs dans la boule de $\mathbb{R}^{r+p_2}$ de rayon δ qui réalise :

$$\overline{\lambda}\left[\ddot{L}(X-\mu_s(X,\delta)\,;\,\sigma_o+\sigma_s(X,\delta)\right] = \sup_{\|(\mu,\sigma)\|\leq\delta} \overline{\lambda}\left[\ddot{L}(X-\mu,\,\sigma_o+\sigma)\right]$$

$$\underline{\lambda}\left[\ddot{L}(X,\mu_i(X,\delta),\sigma_o+\sigma_i(X,\delta)\right] = \inf_{\|(\mu,\sigma)\|\leq\delta} \underline{\lambda}\left[\ddot{L}(X-\mu,\sigma_o+\sigma)\right] .$$

L'hypothèse E combinée avec le théorème de convergence dominée permet d'exhiber δ_1 tel que :

$$E_{\sigma_o} \ddot{L}(X-\mu_s(X,\delta_1),\ \sigma_o+\sigma_s(X,\delta_1)) \leqslant -(1-\frac{\varepsilon'}{2})\ \mathcal{J}(\sigma_o)$$

et $$E_{\sigma_o} \ddot{L}(X-\mu_i(X,\delta_1),\ \sigma_o+\sigma_i(X,\delta_1)) \geqslant -(1+\frac{\varepsilon'}{2})\ \mathcal{J}(\sigma_o)$$

La loi des grands nombres permet alors d'énoncer le :

Lemme (4.7)

Sous les hypothèses A et E, on a :

$\forall \varepsilon > 0$, $\forall \varepsilon' > 0$, $\exists \delta_1(\varepsilon')$ et $\exists k_o(\varepsilon,\varepsilon')$ tels que $\forall n \geq k_o$,

$$P\{\forall k,\ k_o \leqslant k \leqslant n,\ \frac{1}{n} \sum_{j=1}^{k} \mathcal{F}'(\tfrac{j}{n})\, \ddot{L}(X_j-\mu_s(X_j),\sigma_o+\sigma_s(X_j))\, \mathcal{F}(\tfrac{j}{n})$$

$$< -(1-\varepsilon') \sum_{j=1}^{k} \mathcal{F}'(\tfrac{j}{n})\, \mathcal{J}(\sigma_o)\, \mathcal{F}(\tfrac{j}{n})\} \geqslant 1-\varepsilon$$

$$P\{\forall k,\ k_o \leqslant k \leqslant n,\ \frac{1}{n} \sum_{j=1}^{k} \mathcal{F}'(\tfrac{j}{n})\, \ddot{L}(X_j-\mu_i(X_j),\sigma_o+\sigma_i(X_j))\, \mathcal{F}(\tfrac{j}{n})$$

$$\geqslant -(1+\varepsilon') \sum_{j=1}^{k} \mathcal{F}'(\tfrac{j}{n})\, \mathcal{J}(\sigma_o)\, \mathcal{F}(\tfrac{j}{n})\} \geqslant 1-\varepsilon$$

(Nous rappelons que les inégalités ci-dessus sont entendues au sens où la différence est semi-définie positive).

Ce lemme permet de majorer (4.5) dès que le δ définissant B_h est inférieur à $\sqrt{\frac{\gamma}{\Gamma}}\,\delta_1$ car sur B_h, $\theta' \dfrac{\overline{F}(\frac{k}{n})}{\frac{k}{n}}\,\theta \leqslant n\,\delta^2$, ce qui implique : $\sup_{s\in\,]0,\frac{k}{n}]} \| \mathcal{F}(s)\theta \| \leqslant \sqrt{n} \cdot \delta_1$. On obtient :

$$P\{ \sup_{k_o(n)\leq k\leq n} \sup_{\theta\in B_h} Z_n^f(\theta,\frac{k}{n}) > A_n \}$$

$$\leq \varepsilon + P\{ \sup_{k_o(n)\leq k\leq n} \sup_{\theta\in B_h} Q(\theta,k) > \log A_n^q \}$$

avec $Q(\theta,k) = \frac{\partial'}{\sqrt{n}} \sum_{j=1}^{k} \mathscr{F}(\frac{j}{n})' \dot{L}(X_j,\sigma_o) - \frac{(1-\varepsilon')}{2}\underline{\lambda}(\sigma_o)\theta' \bar{F}(\frac{k}{n})\theta$.

A k fixé, le suprémum sur $\mathbb{R}^p$ de $Q(\theta,k)$ est atteint en :

$$\tilde{\theta}_k = \frac{1}{(1-\varepsilon')\underline{\lambda}(\sigma_o)} \bar{F}(\frac{k}{n})^{-1} \frac{1}{\sqrt{n}} \sum_{j=1}^{k} \mathscr{F}(\frac{j}{n})' \dot{L}(X_j,\sigma_o)$$

la majoration se poursuit alors par :

$$\varepsilon + P\{ \bigcup_{k=k_o'(n)}^{n} \tilde{\theta}_k' \bar{F}(\frac{k}{n})\tilde{\theta}_k > \frac{M\frac{k}{n}}{h(\frac{k}{n})} \cap Q(\tilde{\theta}_k,k) > \log A_n^q \}$$

$$+ P\{ \bigcup_{k=k_o'(n)}^{K(n)} \sup \frac{2}{(1-\varepsilon)\underline{\lambda}(\sigma_o)} \log A_n^q < \theta\bar{F}(\frac{k}{n})\theta \; Q(\theta,k) > \log A_n^q \cap \tilde{\theta}_k'\bar{F}(\frac{k}{n})\tilde{\theta}_k < \frac{M\frac{k}{n}}{h(\frac{k}{n})} \}$$

$$+P\{ \sup_{K(n)<k<n} \sup \frac{M\frac{k}{n}}{h(\frac{k}{n})} < \theta\bar{F}(\frac{k}{n})\theta \; Q(\theta,k) > \log A_n^q \}$$

que l'on majore encore en notant

$$Q_n(k) = \frac{1}{n}\Big(\sum_{i=1}^{k} \dot{L}(X_i,\sigma_o)' \mathscr{F}(\frac{i}{n}) \Big) \bar{F}(\frac{k}{n})^{-1} \Big(\sum_{i=1}^{k} \mathscr{F}(\frac{i}{n})' \dot{L}(X_i,\sigma_o) \Big)$$

par :

$$\varepsilon + 2P\{ \bigcup_{k=k_o'(n)}^{K(n)} Q_n(k) > 2(1-\varepsilon)\,\underline{\lambda}(\sigma_o) \log A_n^q \}$$

$$+ 2P\{ \bigcup_{k=K(n)}^{n} Q_n(k) > (1-\varepsilon)^2 \underline{\lambda}^2(\sigma_o) \frac{M\frac{k}{n}}{h(\frac{k}{n})} \}$$

ce qui s'écrit encore :

$$\varepsilon + 2P\Big(\bigcup_{k=k_o'(n)}^{K(n)} \Big[\sum_{i=1}^{k} \dot{L}(X_i,\sigma_o)' \mathcal{F}(\tfrac{i}{n})\Big](A_1(\tfrac{k}{n}))^{-1}\Big[\sum_{i=1}^{k} \mathcal{F}(\tfrac{i}{n})' \dot{L}(X_i,\sigma_o)\Big] \geq 1\Big|$$

$$+ 2P\Big(\bigcup_{k=K(n)}^{n} \Big[\sum_{i=1}^{k} \dot{L}(X_i,\sigma_o)' \mathcal{F}(\tfrac{i}{n})\Big|(A_2(\tfrac{k}{n}))^{-1}\Big[\sum_{i=1}^{k} \mathcal{F}(\tfrac{i}{n})' \dot{L}(X_i,\sigma_o)\Big| \geq 1\Big| \quad (4.4)$$

avec $A_1(\frac{k}{n}) = 2(1-\varepsilon)\underline{\lambda}(\sigma_o)(\log A_n^q)n\overline{F}(\frac{k}{n})$, $A_2(\frac{k}{n}) = (1-\varepsilon)^2\,\underline{\lambda}^2(\sigma_o)\,\dfrac{k\overline{MF}(\frac{k}{n})}{h(\frac{k}{n})}$

Pour achever la démonstration, il convient d'utiliser une inégalité de type maximal donnée par Birnbaum-Marshall [5] que nous énonçons par le :

Lemme (4.8)

Si $(Y_k)_{k=1,\ldots,n}$ est une suite de variables aléatoires adaptées par rapport à une suite de tribus emboitées $(\mathcal{G}_k)$ et satisfait :

$$E(|Y_k| \;\big|\; \mathcal{G}_{k-1}) \geq \psi_k |Y_{k-1}| \qquad \text{p.s.}$$

avec des constantes ψ_k comprises entre 0 et 1 (par convention, $\psi_{n+1} = 0$), alors :

$$P\{\max_{m\leq k\leq n} |Y_k| \geq 1\} \leq \sum_{k=m}^{n} (1-\psi_{k+1})\, E|Y_k|$$

Nous l'appliquons ici pour $j = 1,2$ avec

$$Y_k^{(j)} = \Big[\sum_{i=1}^{k} \dot{L}(X_i,\sigma_o)' \mathcal{F}(\tfrac{i}{n})\Big] A_j(\tfrac{k}{n})^{-1} \Big[\sum_{i=1}^{k} \mathcal{F}(\tfrac{i}{n})' L(X_i,\sigma_o)\Big]$$

$$\mathcal{G}_k = \mathcal{B}(X_1,\ldots,X_k)$$

et $\psi_k^{(j)}$ défini par :

$$E(Y_k \mid \mathcal{G}_{k-1}) \geq \left[\sum_{i=1}^{k-1} \dot{L}(X_i,\sigma_o)' \mathcal{F}(\tfrac{i}{n})\right] A_j(\tfrac{k}{n})^{-1} \left[\sum_{i=1}^{k-1} \mathcal{F}(\tfrac{i}{n})' \dot{L}(X_i,\sigma_o)\right]$$

$$\geq \frac{1}{\|A_j(\frac{k}{n})A_j(\frac{k-1}{n})^{-1}\|} \; Y_{k-1}^{(j)}$$

soit encore : $1 - \psi_k^{(j)} \leq \frac{1}{n} \sup_{\tau \in]\frac{k-1}{n}, \frac{k}{n}[} \|A_j(\tau)A_j(\tau)^{-1}\|$

avec $\dot{A}_j(\cdot)$ désignant la dérivée de la matrice A_j .

Par ailleurs, un calcul simple en utilisant les traces fournit :

$$E|Y_k^{(j)}| < \frac{p\,\|\mathcal{J}(\sigma_o)\|}{2(1-\varepsilon)\underline{\lambda}(\sigma_o)} \left[\frac{1}{\log A_n^q} \cdot 1_{j=1} + \frac{2}{(1-\varepsilon)\underline{\lambda}(\sigma_o)M} \cdot \frac{\frac{k}{n}}{h(\frac{k}{n})} \cdot 1_{j=2}\right]$$

ce qui permet d'achever la démonstration en majorant (4.4) par :

$$\varepsilon + \frac{p\,\|\mathcal{J}(\sigma_o)\|}{(1-\varepsilon)\underline{\lambda}(\sigma_o)} \left| \int_0^{\frac{K(n)}{n}} \|\dot{A}_1(t)A_1(t)^{-1}\| \,(\log A_n^q)^{-1} dt \; + \right.$$

$$\left. \int_{\frac{K(n)}{n}}^1 \frac{2}{(1-\varepsilon)\underline{\lambda}(\sigma_o)M} \|\dot{A}_2(t)\, A_2(t)^{-1}\| \frac{h(t)}{t}\, dt \right.$$

et en exploitant le comportement polynomial de $\overline{F}$ au voisinage de 0, en particulier :

$$\sup_t t\|\mathcal{F}(t)' \mathcal{F}(t)\, \overline{F}(t)^{-1}\| < \sup_t t.\mathrm{tr}|\mathcal{F}(t)\, \overline{F}(t)^{-1}\, \mathcal{F}(t)| < \infty$$

Démonstration de la proposition (1.1)

Elle tient en deux points importants :

i) la convergence étroite, pour tout ε strictement positif, de la restriction à $[\varepsilon,1]$ de la suite de processus :

$$S_n(t) = \sup_{\theta \in \mathbb{R}^p} Z_n^f(\theta,t)$$

ii) la tension en 0 de la suite $\tau_n(t) = \varphi(t) \log S_n(t)$.

Démontrons i)

Pour tous ε et ε' positifs fixés, le lemme (2.1) appliqué avec $\phi(t) = I_{p,p} \cdot 1_{[\varepsilon,1]}(t)$ permet de trouver M satisfaisant pour tout n :

$$P\{\sup_{t\in[\varepsilon,1]} \left| \sup_{\theta,\|\theta\|<M} Z_n^f(\theta,t) - \sup_{\theta,\theta\in\mathbb{R}^p} Z_n^f(\theta,t)\right| = 0\} \geq 1-\varepsilon'$$

et le théorème (1.1) contient la convergence uniforme sur les compacts et donc implique la convergence dans $\mathscr{C}[\varepsilon,1]$ de la suite $S_n(\cdot)$.

Démontrons ii)

Pour ε' fixé, il nous faut trouver ε positif rendant aussi petit que l'on veut :

$$P\{\sup_{t\in[0,\varepsilon]} \varphi(t) \log S_n(t) > \varepsilon'\}$$

Or, pour M quelconque et δ inférieur à $\sqrt{\frac{\gamma}{\Gamma}}\,\delta_0$ (hypothèse E), on a :

$$P\{\sup_{t\in[0,\varepsilon]} \varphi(t) \log S_n(t) > \varepsilon'\}$$

$$\leq P\{\sup_{t\in[0,\varepsilon]} \sup_{\theta,\|\theta\|<M} \varphi(t) \log Z_n^f(\theta,t) > \varepsilon'\}$$

$$+ P\{\sup_{1\leq k\leq n\varepsilon} \sup_{\substack{\theta,\|\theta\|>M \\ \theta'\frac{\bar{F}(\frac{k}{n})}{\frac{k}{n}}\theta \leq n\delta^2}} \varphi(\tfrac{k}{n}) \log Z_n^f(\theta;\tfrac{k}{n}) > \varepsilon'\}$$

$$+ P\{\sup_{1\leq k\leq n\varepsilon} \sup_{\substack{\theta,\|\theta\|>M \\ \theta'\frac{\bar{F}(\frac{k}{n})}{\frac{k}{n}}\theta \geq n\delta^2}} \varphi(\tfrac{k}{n}) \log Z_n^f(\theta;\tfrac{k}{n}) > \varepsilon'\} = \pi_1 + \pi_2 + \pi_3$$

Le premier terme π_1 est tel que pour tous ε' et M, on peut trouver $\varepsilon(\varepsilon',M)$ réalisant pour tout n :

$$\pi_1 \leqslant P\{ \sup_{t\varepsilon[0,\varepsilon]} \sup_{\theta, \|\theta\|<M} Z_n^f(\theta,t) > e^{\frac{\varepsilon'}{\|\varphi\|}} \} \leqslant \varepsilon'$$

à cause de la tension de la suite Z_n^f dans $\mathscr{C}(\mathbb{R}^p \times [0,1])$.

Le second terme π_2 se majore par une technique analogue à la démonstration du théorème (1.1) : si $\ell(n)$ est une suite tendant vers l'infini :

$$P\{ \sup_{1\leqslant k\leqslant n\varepsilon} \sup_{\theta, \|\theta\|>M} \varphi(\tfrac{k}{n}) \log Z_n^f(\theta ; \tfrac{k}{n}) > \varepsilon' \}$$
$$\theta' \frac{\bar{F}(\frac{k}{n})}{\frac{k}{n}} \theta \leqslant n\delta^2$$

$$< P\{ \exists k \in \{\ell(n), n\varepsilon\}, \sup_{\theta, \|\theta\|>M} \frac{\theta'}{\sqrt{n}} \sum_{i=1}^{k} \mathscr{I}(\tfrac{i}{n})' \dot{L}(X_i,\sigma_o)$$
$$\theta' \frac{\bar{F}(\frac{k}{n})}{\frac{k}{n}} \theta \leqslant n\delta^2$$

$$- \frac{1-\varepsilon'}{2} \alpha\, \theta' \bar{F}(\tfrac{k}{n})\theta > \frac{q.\varepsilon'}{\varphi(\frac{k}{n})} \}$$

$$+ P\{ \exists k \in \{1,\ldots,\ell(n)\}, \sup_{\theta\in\mathbb{R}^p} Z_n^f(\theta,\tfrac{k}{n}) > e^{\frac{\varepsilon'}{\varphi(\frac{k}{n})}} \} \qquad (4.7)$$

La première partie se majore par une méthode analogue au lemme (4.6) et la seconde à l'aide de l'hypothèse (H).

Le troisième terme π_3 peut être réduit à cause de la majoration (4.7) à :

$$P\{\exists k \in \{\ell(n),\dots,n\varepsilon\}, \sup_{\theta,\ \theta' \frac{\overline{F}(\frac{k}{n})}{\frac{k}{n}}\theta \geqslant n\delta^2} Z_n^f(\theta,\tfrac{k}{n}) > e^{\frac{\varepsilon'}{\varphi(\frac{k}{n})}}\}$$

$$< \sum_{k=\ell(n)}^{n\varepsilon} B'' \cdot e^{-\frac{q'.\varepsilon'}{\varphi(\frac{k}{n})}} \frac{1}{\left[\det \frac{\overline{F}(\frac{k}{n})}{\frac{k}{n}}\right]^{1/2} (k\,\delta^2)^{N-a'-\frac{p}{2}}} \quad \text{par (4.4)}$$

ce terme est inférieur à ε' pour n assez grand dès que $\ell(n)$ tend vers l'infini et que $\phi(t) < \frac{C'}{(\log \frac{1}{t})^{\frac{s}{q'\varepsilon'}}}$, ce qui est réalisé pour $\varphi(t) \leqslant C.t^{\beta}$ avec $\beta > 0$.

BIBLIOGRAPHIE

[1] P.K. BHATTACHARYA : "Estimation of a small change in the distribution of independant random variables" (1977) Abstract ## 159-40 IMS. bull 6.210.

[2] P.K. BHATTACHARYA - P.J. BROCKWELL : "The minimum of an additive process with applications to signal estimation and storage theory" Z.W.u.v.G. 37, 51-75.

[3] P.K. BHATTACHARYA : Preprint (1981).

[4] P.K. BHATTACHARYA - R.A. JOHNSON : "Non parametric tests for shifts at unknown time points" A.M.S. (1968) n° 39, p. 1731-1743.

[5] Z.W. BIRNBAUM - A.W. MARSHALL : "Some multivariate Chebyshev inequalities with extensions to continuous parameter processes" A.M.S. n° 32 (1961), p. 687-703.

[6] D.M. HAWKINS : "Testing a sequence of observations for a shif in location" J.A.S.A. (1977) n° 72, p. 180-186.

[7] D.V. HINKLEY : "Inference about the change point in a sequence of random variables" Biometrika (1970) n° 57, p. 1.17.

[8] D.V. HINKLEY - E.A. HINKLEY : "Inference about the change point in a sequence of binomial variables" (1970) Biometrika, n° 57 p. 477-488

[9] I.A. IBRAGIMOV - R.Z. KHASMINSKII : "Statistical estimation - Asymptotic theory" Springer verlag - Applications of mathematics, n° 16.

[10] I.A. IBRAGIMOV - R.Z. KASMINSKII : "Local asymptotic normality for non-identically distributed observations" T.P.A. (1975) p. 246-260.

[11] L. LECAM : "Limits of experiments". Proceedings of the VI$^{th.}$ Berkeley Symposium. I. (1972), p. 245-261.

[12] I.B. Mc NEIL : "Tests for change of parameter at unknown time and distributions of some related functionals on Brownian motion" A.S. (1974), n° 2, p. 950-962.

[13] A.N. PETTITT : "A simple cumulative sum type statistic for the change-point problem with zero one observations" Biometrika (1980) n° 67, p. 79-84.

[14] R. REBOLLEDO : "Martingales et convergence étroite des mesures de probabilité". Thèse de doctorat d'Etat (1979) Paris VI.

[15] ROUSSAS : "Contiguity of probability measures : some applications in stotistics" Cambridge University Press.

[16] P.K. SEN : "Asymptotic theory of some tests for constancy of regression relationships over time" (à paraître).

[17] L.A. SHEPP : "The joint density of the maximum and its location for a Wiener process with drift" J.A.P. (1979), n° 16, p. 423-427.

[18] J. DESHAYES - D. PICARD : "Tests de rupture de régression : comparaison asymptotique" Teoryia Ver. Prim. (1982), p. 95-108.

[19] J. DESHAYES - D. PICARD : "Testing for a change-point in statistical models" Prépublications d'Orsay (1980), t. 52.

[20] J. DESHAYES - D. PICARD : "Asymptotique des tests en rupture de modèle" (à paraître).

[21] J. DESHAYES - D. PICARD : "Principe d'invariance sur les processus de vraisemblance" Prépublication d'Orsay (1982), t. 10.

[22] C. DELLACHERIE - P.A. MEYER : "Probabilité et potentiel - théorie des martingales" éd. Hermann.

RUPTURE DE MODELES POUR DES PROCESSUS PONCTUELS.

Jean DESHAYES

Le problème de changement de paramètre (ou de loi) a été étudié dans [14], [15], pour une suite de variables indicées chronologiquement : n-échantillon, modèle de régression, ...

Nous nous posons ici les mêmes questions pour des processus ponctuels : processus de comptage sur la droite. L'article de MATTHEWS-FAREWELL [29] , fournit un exemple sur les temps de rechute après thérapie pour les malades atteints de leucémie, où un modèle de processus de Poisson stationnaire avec un point de rupture semble adapté.

Dans une première partie, nous traiterons les problèmes d'estimation et de tests pour un processus de Poisson homogène. Nous ferons une comparaison des tests classiques pour des tailles d'échantillons fixées puis nous montrerons l'optimalité asymptotique au sens non local du test du rapport de vraisemblance parmi les tests conditionnels ; la comparaison au sens local conduira à l'équivalence des différents tests et montrera, comme dans le problème du n-échantillon, la nécessité de renormaliser certaines statistiques de test. Nous exhiberons la loi limite des estimateurs pour un modèle avec une seule rupture (le cas de plusieurs s'en déduisant immédiatement) pour différents cadres asymptotiques.

Dans une seconde partie, nous traiterons le problème de rupture pour les processus de Poisson non homogènes d'intensité paramétrique : nous ferons une étude asymptotique (au sens local) des méthodes de vraisemblance ici dans un cadre de non stationnarité. Pour une intensité générale, se posera le problème de la non

liberté asymptotique de la statistique de test mais nous donnerons des cas particuliers où elle est libre et donc où il est possible d'approximer le niveau du test. Les lois limites des estimateurs sont semblables à celles du cas stationnaire.

Enfin, dans une troisième partie, nous envisageons des modèles plus généraux où les résultats précédents s'appliquent : processus de Poisson spatial homogène, processus de Poisson composé, processus de vie et de mort, processus "self exciting", processus où les vraisemblances se manipulent bien. Des ruptures de caractéristiques spectrales par des méthodes paramétriques ou non sont également envisagées.

I. RUPTURE D'INTENSITE POUR UN PROCESSUS DE POISSON STATIONNAIRE.

On suppose observer un processus de Poisson N sur un intervalle de temps $[0,T]$ d'intensité connue à un paramètre multiplicatif près : $\alpha(t,\theta) = \theta.f(t)$ où f est une fonction positive connue définie sur $[0,T]$ et θ un paramètre positif inconnu. Par exemple (SNYDER [40] p. 73), la sortie d'un photodétecteur soumis à une source lumineuse peut être modélisé ainsi.

Un changement d'échelle de temps permet dans ce cas de se ramener à un processus de Poisson stationnaire :

$$\text{si } F(t) = \int_0^t f(u)\,du$$

le processus défini par $\tilde{N}(u) = N(F^{-1}(u))$ est un processus de Poisson homogène d'intensité θ sur $[0, F(T)]$. Nous nous posons le problème d'un changement de paramètre θ au cours du temps.

Ce point de vue diffère de celui où on s'intéresse aux intervalles entre les sauts qui suivent des lois exponentielles et où on désire savoir si à partir d'un certain nombre de sauts, le paramètre de la loi exponentielle a changé : dans ce dernier modèle, l'instant de rupture n'est pas le même pour deux échantillons différents ; nous y reviendrons lors de la comparaison des tests.

I.1. Les tests envisagés.

En connaissant le processus $N(.)$ sur $[0,T]$, on doit tester l'hypothèse H_0 : l'intensité est une constante θ sur $[0,T[$, θ positive inconnue - contre l'alternative H_1 : l'intensité vaut θ^* sur $[0,\tau[$ et θ^{**} sur $[\tau,T[$ avec les trois paramètres $\theta^* \in \mathbb{R}^+$, $\theta^{**} \in \mathbb{R}^+$, $\tau \in]0,T[$ inconnus.

Remarque : les sous-alternatives H_1' : $\theta^* > \theta^{**}$ et H'' : $\theta^* < \theta^{**}$ seraient traitées de façon identique.

Si on se restreint aux tests sans biais ayant une fonction puissance continue, définis par une fonctionnelle ϕ définie sur $\mathcal{D}[0,T]$: espace des fonctions continues à droite et ayant une limite à gauche sur $[0,T]$, ces tests seront semblables de niveau a

$$E_{H_0(\theta)}\ \phi\ (N(.)) = a \quad , \quad \forall\theta > 0$$

La statistique $N(T)$ étant exhaustive complète pour le paramètre θ sous H_0, on peut se contenter d'étudier les tests ayant une structure de Neyman c'est-à-dire les tests conditionnellement de niveau a :

$$E_{H_0}\ [\phi\ (N(.)) \mid N(T)] = a$$

D'une façon générale, conditionnellement à $N(T) = n$, le processus de Poisson $N(.)$ d'intensité $\alpha(s)$ sur $[0,T]$ est distribué comme un n-échantillon de densité : $\dfrac{\alpha(s)}{\int_0^T \alpha(u)\,du}$ sur $[0,T]$. Donc, conditionnellement à $N(T) = n$, $\frac{N(.}{n}$ est équivalent en loi, sous $H_1\ (\theta^*,\theta^{**},\tau)$ à la fonction de répartition empiriq d'un n-échantillon de densité :

$$\frac{\theta^*}{\tau\theta^* + (T-\tau)\theta^{**}}\ 1_{[0,\tau[}\ (s) + \frac{\theta^{**}}{\tau\theta^* + (T-\tau)\theta^{**}}\ 1_{[\tau,T[}(s)$$

En fait, cette densité ne dépend que de deux paramètres : par exemple $\tau \in]0,T[$ et le rapport des intensités : $\rho = \frac{\theta^{**}}{\theta^*} \in]0,\infty[$ et conditionnellement à $N(T)=n$, on est ramené au problème de tester l'hypothèse simple :

$\overline{H_0}$: $\rho=1$ (ou $\tau = 0$ ou $\tau = T$) c'est-à-dire qu'on a un n-échantillon de densité uniforme sur $[0,T]$: $\frac{1}{T}\ 1_{[0,T[}(s)$ contre l'hypothèse composée :

$\overline{H_1}$: $\tau \in]0,T[$ et $\rho \in \{]0,1[\ \cup\ ]1,\infty[\}$,c'est-à-dire qu'on a un n-échantillon de densité : $\frac{1}{\tau + (T-\tau)\rho} \cdot 1_{[0,\tau[}(s) + \frac{\rho}{\tau + (T-\tau)\rho}\ 1_{[\tau,T[}(s)$.

Pour tester une hypothèse simple contre une alternative composée qui n'est ni à vraisemblance monotone (elle l'est si τ est connu) ni réductible par

invariance, on va étudier deux tests classiques : le test de Kolmogorov-Smirnov et le test du rapport de vraisemblance.

I.1.a) Le test de Kolmogorov-Smirnov.

Conditionnellement à $N(T) = n$, la statistique de test est :

$$\sup_{0\leq s\leq T} \left| \frac{N(s)}{n} - \frac{s}{T} \right|$$

Il faut la normaliser pour obtenir une loi limite non dégénérée par $\sqrt{n}$, ce qui conduit à la région de rejet de H_0 :

$$\mathcal{D}_1 = \{ \sup_{0\leq s\leq T} \frac{|N(s) - \frac{s}{T} N(T)|}{\sqrt{N(T)}} \geq C_{N(T)}(a) \}$$

$C_n(a)$ étant le seuil du test de Komogorov-Smirnov au niveau a pour un échantillon de taille n.

Remarque : Ce test peut aussi être vu comme le test normalisé basé sur la différence des estimateurs de l'intensité car

$$\frac{N(s) - \frac{s}{T} N(T)}{\sqrt{N(T)}} = \frac{s\ (T-s)}{T\ \sqrt{N(T)}} \left(\frac{N(s)}{s} - \frac{N(T)-N(s)}{T-s} \right)$$

I.1.b) Le test du rapport de vraisemblance.

Conditionnellement à $N(T) = n$, le rapport de vraisemblance vaut :

$$\left[\frac{T}{\tau + \rho(T-\tau)} \right]^n \rho^{n-N(\tau)}$$

A τ fixé, le suprémum est atteint en $\hat{\rho}_\tau = \frac{n-N(\tau)}{T-\tau} \cdot \frac{\tau}{N(\tau)}$ et le log. du suprémum vaut :

$$N(T)\ .\ K\left(\frac{N(\tau)}{N(T)}\ ,\ \frac{\tau}{T}\right)$$

où $K(p,q)$ désigne l'information de Kullback : $p \operatorname{Log} \frac{p}{q} + (1-p) \operatorname{Log} \frac{1-p}{1-q}$

ce qui amène à la région de rejet de H_0 :

$$\mathcal{D}_2 = \{ \sup_{0\leq s\leq T} N(T)\ \ K\left[\frac{N(s)}{N(T)}\ ,\ \frac{s}{T}\right] \geq d_{N(T)}\ (a) \}$$

$d_n(a)$ étant le seuil du test de niveau a pour un n-échantillon de loi uniforme sur $[0,1]$:

$$P\{n.\sup_{0\leq u\leq 1} K[F_n(u), u] \geq d_n(a)\} = a \quad .$$

Ce dernier test est aussi un test non paramétrique, il a d'ailleurs été étudié par BERK-JONES [7] comme alternative au test de Kolmogorov-Smirnov.

Nous verrons au § I.3 que la statistique de test ainsi définie ne converge pas sous $\overline{H}_o$, il est nécessaire de pondérer par une fonction $\varphi(s)$ avant de prendre le suprémum ce qui revient à faire franchir une frontière $h(.)$ à la marche aléatoire considérée :

$$\mathscr{D}'_2 = \{ \exists s \in]0,T[, \quad N(t) \quad K\left(\frac{N(s)}{N(T)}, \frac{s}{T}\right) \geq h(s)\}$$

A condition de choisir des frontières supérieures et inférieures différentes pour la statistique de Kolmogorov-Smirnov, on obtient des tests équivalents.

PROPOSITION I.1. :

A toute barrière $h(.)$ définissant une région de rejet du test du rapport de vraisemblance, on peut faire correspondre deux frontières $\underline{h}(.)$ et $\overline{h}(.)$ telles que les régions $\mathscr{D}'_1$ et $\mathscr{D}'_2$ soient confondues :

$$\mathscr{D}'_1 = \{\exists\ s \in]0,T[\ , \frac{N(s) - \frac{s}{T} N(T)}{\sqrt{N(T)}} \leq \underline{h}(s) \quad \text{ou} \quad \frac{N(s) - \frac{s}{T} N(T)}{\sqrt{N(T)}} \geq \overline{h}(s)\}$$

Démonstration : elle repose sur la convexité par rapport à la variable p de l'information de Kullback : $p \operatorname{Log} \frac{p}{q} + (1-p) \operatorname{Log} \frac{1-p}{1-q}$. Toutefois il n'est pas possible d'écrire explicitement les barrières $\underline{h}(.)$ et $\overline{h}(.)$; on y parviendra asymptotiquement (n tendant vers l'infini) au paragraphe I.3.

On se propose maintenant de déterminer les meilleures barrières et de comparer aussi les tests $\mathscr{D}_1$ et $\mathscr{D}_2$ correspondants aux barrières fixes par une étude asymptotique au sens non local (§ I.2) et au sens local (§ I.3).

Pour ce faire, deux possibilités se présentent : faire tendre l'intensité vers l'infini sur un intervalle de longueur fixe, ce qui correspond à une superposition de processus de Poisson indépendants, identique distribués, ou bien faire tendre la longueur de l'intervalle d'observation vers l'infini ; ces deux procédés sont strictement équivalents car par un changement de temps dans la seconde situation, on se ramène à la première :

sous $H_{1,T}$: l'intensité du processus $N_T(.)$ vaut $\theta^*.1_{[0,\tau_T[}(.)+\theta^{**}.1_{[\tau_T,T[}(.)$ l'intervalle $[0,T]$ devenant infini, on suppose que l'instant de rupture τ_T est tel que : τ_T et $T-\tau_T$ tendent vers l'infini. Le nouveau processus $\tilde{N}(.)$ défini sur $[0,1]$ par

$$\tilde{N}(s) = N_T(sT)$$

nous ramène à la première situation. Nous choisissons donc le cadre asymptotique suivant :

$N_\nu(.)$ est une suite de processus de Poisson sur $[0,1]$ d'intensités :

θ_ν sous $H_{0,\nu}$ avec $\theta_\nu \xrightarrow[\nu\to\infty]{} \infty$.

$\theta^*_\nu \ 1_{[0,\tau_\nu[} + \theta^{**}_\nu \ 1_{[\tau_\nu,1[}$ sous $H_{1,\nu}$ avec $\theta^*_\nu . \tau_\nu$ et $\theta^{**}_\nu(1-\tau_\nu) \xrightarrow[\nu\to\infty]{} \infty$.

I.2. *Comparaison des tests au sens des grandes déviations.*

Lorsque ν tend vers l'infini, $N_\nu(1)$ tend vers $+\infty$ presque sûrement sous $H_{0,\nu}$ et sous $H_{1,\nu}$; les tests conditionnels à $N_\nu(1)=n$ définis précédemment vont donc être étudiés lorsque ν et n tendent vers l'infini.

Conditionnellement à $N_\nu(1)=n$, la vitesse naturelle de décroissance des erreurs est $\frac{1}{n}$. Pour le test du rapport de vraisemblance , on a :

$$\lim_{n\to\infty} \frac{1}{n} \text{Log } P_{H_{0,\nu}} \{ \sup_{0\leq s\leq 1} K[\frac{N_\nu(s)}{N_\nu(1)}, s] \geq d \,|N_\nu(1) = n\}$$

$$= \lim_{n\to\infty} \frac{1}{n} \text{Log } P_{\bar{H}_{0,n}} \{ \sup_{0\leq u\leq 1} K[F_n(u),u] \geq d \} = - d$$

d'après les calculs de grandes déviations pour mesures empiriques ([41]) .

D'autre part, conditionnellement à $N_\nu(1)=n$, nous obtenons la région la plus grande de niveau exponentiel : -d car le franchissement de la barrière pour n'importe quelle valeur de s comprise entre 0 et 1 fournit le bon niveau exponentiel : -d ,ce qui permet d'affirmer l'optimalité asymptotique au sens de BROWN [10

PROPOSITION 1.2 :

Le test du rapport de vraisemblance est optimal au sens des vitesses de décroissance exponentielle des erreurs parmi les tests conditionnels en $N_\nu(1)$.

Remarques :

i) L'optimalité au sens de Bahadur du test non paramétrique défini par la région :

$$\{\sup_{0<u<1} K[F_n(u),u] \geq d\}$$

est prouvée dans BERK-JONES [7] pour de vastes contre-hypothèses incluant $\bar{H}_1$; une optimalité analogue pour d'autres tests conditionnels a été prouvée par BASAWA [5].

ii) Le niveau conditionnel satisfait aussi la relation de convergence :

$$\lim_{\nu\to\infty} \frac{1}{\theta_\nu} \text{Log } P_{H_{0,\nu}} \{ \sup_{0\leq s\leq 1} K[\frac{N_\nu(s)}{N_\nu(1)}, s] \geq d \,|N_\nu(1)\} = - d$$

en $P_{H_0(\theta_\nu)}$- probabilité .

On peut aussi vérifier par un calcul direct de grandes déviations pour marches aléatoires que le niveau non conditionnel vaut aussi -d à la limite :

$$\lim_{\nu\to\infty} \frac{1}{\theta_\nu} \operatorname{Log} P_{H_o(\theta_\nu)} \{ \sup_{0\leq s\leq 1} K [\frac{N_\nu(s)}{N_\nu(1)}, s] \geq d \} = -d$$

Cette optimalité au sens des grandes déviations du test de vraisemblance parmi les tests conditionnels a malheureusement peu de retombées sur le plan pratique ; en effet, comme dans le cas d'un n-échantillon [15] , la recherche d'un seuil $d_n(\alpha)$ fournissant un niveau α fixé à l'avance conduit à une suite $d_n(\alpha)$ tendant vers l'infini lorsque n tend vers l'infini. Ceci est dû à la non convergence sous H_o de la statistique de suprémum aux bords de $[0,1]$.

I.3. Comparaison des tests au sens local.

Comme précédemment, nous pouvons étudier les deux tests conditionnellement à $N_\nu(1) = n$ et faire tendre n vers l'infini ; le comportement des statistiques de test non conditionnellement lorsque ν tend vers l'infini et le comportement des statistiques conditionnellement à $N_\nu(1) = n$, lorsque n tend vers l'infini se déduisent l'un de l'autre car $N_\nu(1)$ tend vers l'infini en probabilité sous $H_o(\theta_\nu)$.

La proposition suivante exhibe les renormalisations possibles de la statistiques du rapport de vraisemblance près de 0 et près de 1 et son équivalence asymptotique avec celle de Kolmogorov-Smirnov.

PROPOSITION (1.3) :

Si $F_n(.)$ désigne la fonction de répartition empirique d'un n-échantillon de loi uniforme sur $[0,1]$, on a

i) la statistique non renormalisée

$$n \sup_{0\leq u\leq 1} K [F_n(u), u]$$

tend vers l'infini lorsque n tend vers l'infini.

ii) Si φ est une fonction positive continue sur $[0,1]$ telle que

$\int_0^1 \frac{\varphi^2(t)}{t(1-t)}\,dt < \infty$ la suite de processus

$$S_n(u) = n\ \varphi(u)\ K\,[F_n(u),\ u]$$

converge étroitement dans $\mathscr{C}[0,1]$ vers

$$S(u) = \varphi(u)\,\frac{B^2(u)}{2u(1-u)}$$

où $B(.)$ est le pont brownien sur $[0,1]$. On a même le résultat plus fort :

$$\text{iii)} \quad \sup_{u\in[0,1]} \; n\,\varphi(u)\,\left|\frac{[F_n(u)-u]^2}{2u(1-u)} - K\,[F_n(u),u]\right|$$

tend vers 0 en probabilité lorsque n tend vers l'infini.

Remarques :

a) Pour se rapprocher du test optimal au sens des grandes déviations, on aurait intérêt à choisir φ la plus petite possible, par exemple $\varphi(u) = \left[\text{Log}\,\frac{1}{u(1-u)}\right]^{\eta}$ avec $\eta < -\frac{1}{2}$; mais le calcul effectif d'un niveau approché du test est facilité avec $\varphi(u) = 2u(1-u)$ d'après ii).

b) L'assertion iii) montre que le test de vraisemblance renormalisé par $\varphi(u) = 2u(1-u)$ est équivalent au test de Kolmogorov asymptotiquement au sens local puisque les deux tests auraont même puissance sous une hypothèse contigue à H_o .

Démonstration :

L'assertion i) sera impliquée par ii) en prenant une suite de fonctions φ_ℓ nulles en 0 et 1, égales à 1 sur $[\frac{1}{\ell},\ 1-\frac{1}{\ell}]$ et linéaires sur $[0,\ \frac{1}{\ell}]$ et $[1-\frac{1}{\ell},\ 1]$; en effet, pour tous ℓ et n

$$n \sup_{0\le u\le 1} K\,[F_n(u),u] \ge \sup_{0\le u\le 1} n\,\varphi_\ell(u)\,K\,(F_n(u),u)$$

à ℓ fixé, le second membre convergeant en loi vers $\sup_{0\le u\le 1} \varphi_\ell(u)\frac{B^2(u)}{2u(1-u)}$ et cette dernières statistique convergeant vers l'infini lorsque ℓ tend vers l'infini (loi du logarithme itéré pour le pont brownien), on en déduit que le premier membre tend vers l'infini en probabilité lorsque n tend vers l'infini.

L'assertion ii) peut être montrée directement sur le processus de Poisson homogène N_ν et la convergence de la suite de processus $S_n(.)$ s'en déduit en décomposant sur les événements $\{N_\nu(1)\} = n$: ce procédé est celui utilisé par PYKE-SHORACK [38] pour montrer la convergence de la suite $S'_n(u) = \sqrt{\frac{n}{2u(1-u)}}\,[F_n(u)-u]$ dans $\mathscr{C}[0,1]$ vers $\sqrt{\frac{\varphi(u)}{2u(1-u)}}\,B(u)$. Nous ferons une démonstration directe de la convergence de la suite $S_n(.)$ au paragraphe II.3 avec une condition un peu plus forte sur φ à cause de la non-stationnarité.

On peut aussi faire une démonstration de ii) et iii) simultanément en utilisant le résultat de [17] :

$$\forall\, \varepsilon > 0, \quad \exists B > 0 \quad \text{tel que} \quad P\left\{ \sup_{t\in[0,1]} \frac{F_n(t)}{t} > B \right\} < \frac{\varepsilon}{3}$$

ceci permet d'assurer la tension de la suite $S_n(.)$ au voisinage de 0 (et aussi de 1) :

$$\forall \varepsilon > 0, \quad \exists \delta(\varepsilon) \text{ et } n_1(\varepsilon) \text{ tels que } \forall n > n_1 ,$$

$$P\left\{ \sup_{0\le u\le \delta(\varepsilon)} S_n(u) \ge \varepsilon \right\} \le \frac{\varepsilon}{3}$$

Pour $\delta(\varepsilon)$ fixé, la proximité des 2 processus $S_n(.)$ et $S'_n(.)^2$ sur $[\delta,1-\delta]$ est vérifiée à l'aide d'un développement de Taylor de la fonction Log à l'ordre 3 :

$$\forall \varepsilon > 0 , \quad \exists A\,(\varepsilon,\delta) \text{ tel que } \forall\, n ,$$

$$P\left\{ \sup_{\delta\le u\le 1-\delta} \sqrt{n}\,|F_n(u)-u| \ge \delta\,(1-\delta)A \right\} \le \frac{\varepsilon}{3}$$

A fortiori, on a :

$$P\left\{ \sup_{\delta\le u\le 1-\delta} \left|\frac{F_n(u)}{u} - 1\right| \le \frac{A}{\sqrt{n}} \;\cup\; \sup_{\delta\le u\le 1-\delta} \left|\frac{1-F_n(u)}{1-u} - 1\right| \le \frac{A}{\sqrt{n}} \right\} \ge 1 - \frac{\varepsilon}{3}$$

Le développement de Taylor de Log fournit une constante $B(\delta,A)$ telle que ce dernier événement est inclus dans : $\{ \sup_{\delta\le u\le 1-\delta} |S'^2_n(u)-S_n(u)| \le \frac{B}{\sqrt{n}} \}$ ce qui implique pour : $n \ge \frac{B^2}{\varepsilon^2}$, $P\left\{ \sup_{\delta\le u\le 1-\delta} |S'^2_n(u)-S_n(u)| \ge \varepsilon \right\} \le \frac{\varepsilon}{3}$ et achève la démonstration de la proposition.

I.4. Loi des estimateurs du maximum de vraisemblance pour un processus de Poisson homogène avec changement de paramètre.

Nous voulons estimer les paramètres τ , θ^* et θ^{**} qui sont respectivement l'instant de rupture et les intensités avant et après rupture d'un processus de Poisson N observé sur $[0,1]$. Pour établir les lois des estimateurs du maximum de vraisemblance, il est commode de faire le changement de paramétrage déjà introduit au I.1°. :

$$\mu = \tau\theta^* + (1-\tau)\,\theta^{**} \quad , \quad \rho = \frac{\theta^{**}}{\theta^*}$$

L'estimateur du maximum de vraisemblance $\hat{\mu}$ de μ est évidemment $\hat{\mu} = N(1)$ et suit donc une loi de Poisson de paramètre μ sous $H_1(\mu,\rho,\tau)$. Pour trouver les lois des estimateurs $\hat{\rho}$ et $\hat{\tau}$, il semble plus facile de les établir d'abord conditionnellement à $N(1)$. Nous savons que conditionnellement à $N(1) = n$, la vraisemblance est celle d'un n-échantillon de densité :

$$\frac{1}{\tau+(1-\tau)\rho}\ 1_{[0,\,\tau[}(s) + \frac{\rho}{\tau+(1-\tau)\rho}\ 1_{[\tau,1[}(s)$$

Pour τ fixé, l'estimateur de ρ est $\hat{\rho}_\tau = \frac{n-N(\tau)}{1-\tau} \cdot \frac{\tau}{N(\tau)}$ et l'estimateur de τ est donc $\hat{\tau}$ défini par l'instant du suprémum de :

$$L(t) = N(t)\ \mathrm{Log}\ \frac{N(t)}{t.n} + [n-N(t)]\ \mathrm{Log}\ \frac{n-N(t)}{(1-t)n}$$

$$= n\ K\,[F_n(t),t]$$

La loi de τ admet une densité par rapport à la mesure de Lebesgue sur $[0,1]$ mais elle dépend des paramètres τ et ρ et ne peut être calculée qu'asymptotiquement.

Il serait possible d'établir les lois limites de $\hat{\tau}$ et de $\hat{\rho}$ lorsque n tend vers l'infini en utilisant le fait que nous avons un n-échantillon dont la loi est régulière par rapport au paramètre ρ mais présente une discontinuité par rapport au paramètre τ ; en suivant IBRAGIMOV-KHASMINSKII [23] (chapitre V),

la convergence du logarithme du processus de vraisemblance vers une combinaison de processus gaussiens et de processus de Poisson donnerait le résultat mais pour le processus de Poisson homogène, nous allons faire une démonstration directe lorsque l'intensité totale μ tend vers l'infini, les autres paramètres τ et ρ restant fixes.

Définissons le processus de vraisemblance au point (μ,ρ,τ) par :

$$\mathbb{R} \times \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}^+$$

$$(m,r,t) \qquad Z_\mu(m,r,t) = \frac{dP_{H_1(\mu + m\sqrt{\mu},\rho+ \frac{r}{\sqrt{\mu}},\tau + \frac{t}{\mu})}}{dP_{H_1(\mu,\rho,\tau)}} (N_\mu)$$

Pour $t > 0$, un calcul immédiat fournit :

$$\text{Log } Z_\mu (m,r,t) = - m\sqrt{\mu} + N_\mu(\tau) \text{ Log } \frac{\mu + m\sqrt{\mu}}{(\tau+ \frac{t}{\mu}) + (\rho+ \frac{r}{\sqrt{\mu}}) (1-\tau- \frac{t}{\mu})} \times \frac{\tau+\rho(1-\tau)}{\mu}$$

$$+ [N_\mu(\tau+ \tfrac{t}{\mu})- N_\mu(\tau)] \text{ Log } \frac{\mu + m\sqrt{\mu}}{(\tau+ \frac{t}{\mu})+(\rho+ \frac{r}{\sqrt{\mu}})(1-\tau- \frac{t}{\mu})} \times \frac{\tau+\rho(1-\tau)}{\rho.\mu}$$

$$+ [N_\mu(1)-N_\mu(\tau+ \tfrac{t}{\mu})] \text{ Log } \frac{(\rho+ \frac{r}{\sqrt{\mu}}) (\mu + m\sqrt{\mu})}{(\tau+ \frac{t}{\mu}) + (\rho+ \frac{r}{\sqrt{\mu}})(1-\tau- \frac{t}{\mu})} \times \frac{\tau+\rho(1-\tau)}{\rho.\mu}$$

Les trajectoires de Z_μ sont continues en m et r et c.a.d.l.a.g. en t et nous allons montrer la convergence de la suite Z_μ dans l'espace $CD_o(\mathbb{R}^2 x \mathbb{R}; \mathbb{R}^+)$ des fonctions continues en (m,r) et c.a.d.l.a.g. en t muni d'une topologie "uniforme-Skorohod" sur les compacts et de convergence uniforme vers 0 à l'infini.

THEOREME 1.5.

La suite de processus Z_μ converge étroitement sous $H_1(\mu,\rho,\tau)$ dans $CD_o(\mathbb{R}^2 x \mathbb{R}; \mathbb{R}^+)$ vers le processus Z défini par :

$$\text{Log } Z(m,r,t) = m.\xi_1 - \frac{m^2}{2} + r.\frac{1}{\tau+(1-\tau)\rho}\sqrt{\frac{\tau(1-\tau)}{\rho}}\,\xi_2 - \frac{r^2\,\tau(1-\tau)}{2[\tau+(1-\tau)]^2\rho}$$

$$+ \text{Log}\,\frac{1}{\rho}\;\bar{N}(t) - \frac{(1-\rho)}{\tau+(1-\tau)\rho}\;.\;t$$

avec $\bar{N}(t)$: Processus de Poisson sur $\mathbb{R}$ d'intensité $\frac{1}{\tau+(1-\tau)\rho}\,1_{]-\infty,0[}(t) +$

$$+\frac{\rho}{\tau+(1-\tau)}\,1_{[0,\infty[}(t)$$

nul en 0

ξ_1 et ξ_2 : variables gaussiennes centrées, réduites indépendantes et indépendantes de N .

COROLLAIRE (1.6) :

Lorsque l'intensité totale μ tend vers l'infini, les paramètres ρ et τ restant fixes, les estimateurs du maximum de vraisemblance de ces trois paramètres sont asymptotiquement indépendants et

$$\frac{\hat{\mu}-\mu}{\sqrt{\mu}} \xrightarrow{\mathcal{L}} \mathcal{N}(0;1)$$

$$\sqrt{\mu}\,(\hat{\rho}-\rho) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0\;;\;\frac{\tau(1-\tau)}{\rho[\tau+(1-\tau)\rho]^2}\right)$$

$$\mu\,\frac{(\hat{\tau}-\tau)}{\tau+(1-\tau)\rho} \xrightarrow{\mathcal{L}} \zeta_\rho$$

ζ_ρ est la loi sur $\mathbb{R}$ de l'instant de suprémum du processus

$$\text{Log}\,\frac{1}{\rho}\;.\;\bar{N}_\rho\,(u) - (1-\rho)\,u$$

$\bar{N}_\rho(.)$ étant le processus de Poisson sur $\mathbb{R}$ d'intensité $1_{]-\infty,0[}(u)+\rho.1_{[0,\infty[}(u)$

Démonstration : L'expression simple du logarithme du processus de vraisemblance en fonction du processus de Poisson N_μ permet d'utiliser directement la convergence en loi de $\left(\frac{1}{\sqrt{\mu}}\,[N_\mu(\tau) - \mu\,\frac{\tau}{\tau+(1-\tau)\rho}\,]\;,\;\frac{1}{\sqrt{\mu}}\,[N_\mu(1)-\mu]\right)$ vers un couple gaussien centré et le fait que la suite de processus $(N_\mu(\tau+\frac{t}{\mu})- N_\mu(\tau))$ est asymptotiquement indépendante du couple précité et de même loi : celle de $\bar{N}$.

La loi asymptotique de $\hat{\mu}$ et de $\hat{\rho}$ peut être calculée directement mais celle de $\hat{\tau}$ pose le problème qu'elle dépend de ρ . Nous nous retrouvons dans une situation analogue à celle de HINKLEY [20],[21] où une rupture d'amplitude fixe Δ du paramètre d'une loi de Bernouilli ou d'une loi gaussienne nécessitait un calcul de loi limite différent pour chaque valeur de Δ.

Pour obtenir une seule loi limite à calculer, il faut être dans un cadre asymptotique différent et supposer que ρ dépend de μ : ρ_μ tendant vers 1 lorsque μ tend vers l'infini. Cela change la normalisation du paramètre τ et remplace le processus limite de Poisson $\overline{N}$ par un mouvement brownien W sur $\mathbb{R}$; la loi asymptotique de $\hat{\tau}$ est alors dérivée comme dans [15] de celle de l'instant de suprémum de : $W(u) - \frac{|u|}{2}$. Nous ferons une démonstration détaillée de ce résultat au II dans un modèle plus général : rupture d'intensité d'un processus de Poisson non homogène.

II. PROCESSUS DE POISSON NON STATIONNAIRE : ASYMPTOTIQUE DES METHODES DE VRAISEMBLANCE.

Nous nous intéressons maintenant à la détection d'un changement de paramètre dans un processus de Poisson (multidimensionnel) d'intensité paramétrique : $\alpha(t,\theta)$ sur un intervalle de temps ramené à $[0,1]$ comme précédemment :

$N = (N^{(1)},\ldots,N^{(r)})$ et $\alpha = (\alpha^{(1)},\ldots,\alpha^{(r)})$ sont supposés de dimension r, $\int_0^1 \alpha^{(j)}(s,\theta)ds < \infty$, et le paramètre θ appartient à un ouvert Θ de $\mathbb{R}^p$. De nombreux exemples de processus de Poisson non homogènes sont traités dans la littérature ; citons-en trois :

Exemple 2.1 : l'émission de rayons Gamma par une source radioactive se modélise ainsi avec une intensité : $\theta^{(1)}.e^{-\theta^{(2)}.t}$, où $\theta^{(1)}$ dépend de la quantité de source au départ et $\theta^{(2)}$ est l'inverse de la vie moyenne de cette source ([40]).

Exemple 2.2 : en médecine nucléaire, les combinaisons des intensités analogues à la précédente se rencontrent fréquemment ([40])

$$\alpha(t,\theta) = \sum_{j=1}^{q} \theta_j^{(1)} e^{-\theta_j^{(2)}t}$$

Exemple 2.3 :En séismologie et en électrophysiologie acoustique,les impulsions étroites de signaux peuvent être modélisées dans le temps avec une intensité de la forme

$$\alpha(t,\theta) = \theta^{(1)} e^{\theta^{(2)}\cos(2\pi ft+\theta^{(3)})} \qquad \text{(cf.[40], [42])}$$

Beaucoup de modèles de Poisson non homogènes répondent en fait à des intensités log-linéaires : $\alpha(t,\theta) = e^{t\theta f(t)}$ où $f(t)$ est une fonction vectorielle connue ; nous ferons des remarques sur ce modèle particulier dans la suite. D'ores et déjà nous pouvons dire que $\int_0^t f(s)\, dN(s)$ forme un résumé exhaustif et la convergence du processus de vraisemblance serait beaucoup plus simple. Pour faire une étude asymptotique, nous ferons comme au premier paragraphe tendre l'inten-

sité vers l'infini sur l'intervalle fixe [0,1]. Cela peut correspondre au fait qu'un des paramètres du modèle devient grand ; pour ne pas le spécifier, nous écrirons une suite de modèles : N_n d'intensités : $\alpha_n(t,\theta) = n.\alpha(t,\theta)$ et ferons tendre n vers l'infini. Nous supposerons que l'intensité $\alpha(.,.)$ est régulière pour établir un théorème de convergence du processus de vraisemblance. Le processus vectoriel $N = (N^{(1)},\ldots,N^{(r)})$ étant supposé d'intensité déterministe et continue, ses coordonnées sont des processus de Poisson indépendants et il suffit d'établir la convergence du processus de vraisemblance pour chaque coordonnée séparément.

Nous supposerons donc aux paragraphes II.1.2.3. que le processus est unidimensionnel : r=1. Nous serons alors en mesure d'en déduire les lois asymptotiques des statistiques de test et des estimateurs du maximum de vraisemblance aux paragraphes II.4 et II.5 où nous reprendrons le cas multidimensionnel.

II.1. *Hypothèses de régularité de l'intensité.*

Le rapport des vraisemblances correspondant aux processus de Poisson non homogènes d'intensités respectives $\alpha(.,\theta)$ et $\alpha(.,\theta_0)$ sur $[0,t]$ vaut :

$$e^{+\int_0^t \text{Log}\,\frac{\alpha(s,\theta)}{\alpha(s,\theta_0)}\,dN_1(s) - \int_0^t [\alpha(s,\theta)-\alpha(s,\theta_0)]ds}$$

d'où l'information de Kullback entre ces deux lois sur $[0,t]$:

$$\int_0^t \text{Log}\,\frac{\alpha(s,\theta_0)}{\alpha(s,\theta)}\,.\alpha(s,\theta_0)ds - \int_0^t [\alpha(s,\theta_0)-\alpha(s,\theta)]\,ds$$

et l'information de Fisher en est la dérivée seconde en θ au point θ_0 :

$$I(\theta_0;[0,t]) = \int_0^t \frac{\frac{\partial}{\partial\theta}\alpha(s,\theta_0)}{\alpha(s,\theta_0)}\;\frac{\frac{\partial}{\partial\theta}\alpha(s,\theta_0)}{\alpha(s,\theta_0)}\,\alpha(s,\theta_0)\,ds = 4\int_0^t \frac{\partial}{\partial\theta}\sqrt{\alpha(s,\theta_0)}.\frac{\partial}{\partial\theta}\sqrt{\alpha(s,\theta_0)}\,ds$$

nous obtenons une information analogue à celle donnée par une variable aléatoire de "densité" $\alpha(.,\theta)$ par rapport à la mesure de Lebesgue .

De façon plus précise, si l'on fait un changement de paramètre de sorte que l'un d'entre eux représente l'intensité totale sur $[0,t]$:

$$\alpha(s,\theta) = m_t(\theta) \cdot \tilde{\alpha}\,(s,\, g_t(\theta))$$

avec $$m_t(\theta) = \int_0^t \alpha(s,\theta)ds \quad \text{et} \quad \int_0^t \tilde{\alpha}(s,\eta)\, ds = 1$$

Ce nouveau paramétrage sur $[0,t]$ fait apparaître une vraie densité : $\tilde{\alpha}$ et montre que les paramètres m et η s'estiment indépendamment :

$$I(m,\eta;[0,t]) = \begin{pmatrix} \frac{1}{m} & 0 \\ 0 & 4m \int_0^t \frac{\partial}{\partial\eta}\, \tilde{\alpha}^{1/2}(s,\eta)\, \frac{t\partial}{\partial\eta}\, \tilde{\alpha}^{1/2}(s,\eta)ds \end{pmatrix}$$

l'explication vient du fait que à η connu, $N(t)$ est une statistique exhaustive pour le pramètre m et que conditionnellement à $N(t)=\nu$, l'observation est identique à celle d'un ν-échantillon de loi admettant la densité :

$$\tilde{\alpha}(s,\eta) \quad \text{sur} \quad [0,t]$$

Nous sommes conduits naturellement à supposer sur l'intensité des hypothèses de régularité identiques à celles posées sur la densité de probabilité d'un n-échantillon (voir KUTOYANTS [27] pour des hypothèses analogues avec un intervalle d'observation $[0,T]$ devenant infini) ; la possibilité d'observer une rupture en un instant quelconque τ de $]0,1[$ conduit de plus à supposer ces hypothèses sur l'intensité restreinte à un sous-intervalle quelconque de $[0,1]$.

$\underline{\underline{A}}$: i) la fonction $\theta \to \alpha(t,\theta)$ est absolument continue sur $\bar{\Theta}$, $\forall t \in [0,1]$.

ii) l'information de Fisher correspondant à tout sous-intervalle $[t_1,t_2]$ de $[0,1]$

$$I(\theta;[t_1,t_2]) = 4 \int_{t_1}^{t_2} \frac{\partial}{\partial\theta}\, \alpha^{1/2}(s,\theta)\, \frac{t\partial}{\partial\theta}\, \alpha^{1/2}\,(s,\theta)ds \;, \quad t_1 < t_2$$

est continue sur $\bar{\Theta}$, et définie positive au voisinage du point θ_0 .

iii) Tout intervalle $]t_1,t_2[\subset [0,1]$ sépare les paramètres au voisinage de θ_o c'est-à-dire que pour tous δ et δ' positifs :

$$\inf_{\theta_1,\ \|\theta_1-\theta_o\|<\delta'} \ \inf_{\theta,\ \|\theta-\theta_1\|>\delta} \int_{t_1}^{t_2} (\sqrt{\alpha(s,\theta)} - \sqrt{\alpha(s,\theta_1)})^2\, ds > 0.$$

B: Il existe des réels $a > 0$ et $q > (p \vee 2)$ tels que

$$\text{i)} \quad \sup_{\theta\in\Theta} (1 \vee \|\theta\|^a) \int_0^1 \frac{\partial}{\partial\theta}\sqrt{\alpha(s,\theta)}\ \frac{\partial}{\partial\theta}\sqrt{\alpha(s,\theta)}\, ds < \infty$$

$$\text{ii)} \quad \sup_{\theta\in\Theta} (1 \vee \|\theta\|^a) \int_0^1 \left\| \frac{\partial}{\partial\theta}\sqrt{\alpha(s,\theta)} \right\|^q \alpha(s,\theta)^{1-\frac{q}{2}}\, ds < \infty$$

C: Il existe des réels δ et δ' positifs tels que pour tous $t_1 < t_2$,

$$\sup_{\theta_1,\ \|\theta_1-\theta_o\|\le\delta'} \ \sup_{\theta\in\Theta-\theta_1} \{t_\theta\, I(\theta_1;[t_1,t_2])\theta\}^\delta\, e^{-\frac{1}{2}\int_{t_1}^{t_2} (\sqrt{\alpha(s,\theta_1+\theta)} - \sqrt{\alpha(s,\theta_1)})^2 ds} \le 1$$

D: Il existe δ_o positif tel que :

$$\text{i)} \quad \sup_{\theta_1,\ \|\theta_1-\theta_o\|\le\delta'} \ \sup_{\theta,\ \|\theta-\theta_1\|\le\delta_o} \int_0^1 \frac{\frac{\partial\alpha}{\partial\theta}(s,\theta)}{\alpha(s,\theta)} \cdot \frac{\frac{\partial\alpha}{\partial\theta}(s,\theta)}{\alpha(s,\theta)} \cdot \alpha(s,\theta_1) ds < \infty$$

$$\text{ii)} \quad \sup_{\theta_1,\ \|\theta_1-\theta_o\|\le\delta'} \ \sup_{\theta,\ \|\theta-\theta_1\|\le\delta_o} \int_0^1 \frac{\left\|\frac{\partial^2\alpha}{\partial\theta^2}(s,\theta)\right\|}{\alpha(s,\theta)} \cdot \alpha(s,\theta_1)\, ds < \infty$$

avec la matrice hessienne $\frac{\partial^2}{\partial\theta^2}\alpha(s,\theta)$ supposée continue en θ.

F: Pour tout θ_1 appartenant à un voisinage de θ_o, il existe des constantes γ et Γ telles que pour tous $t_1 < t_2$, il existe $t_3 < t_4$ tous inclus dans $[0,1]$ satisfaisant :

$$[t_3,t_4] \subset [t_1,t_2] \quad \text{et} \quad t_4-t_3 \ge \gamma(t_2-t_1)$$

et $\forall u \in \mathbb{R}^p$, $\sup_{s\in[t_1,t_2]} |{}^t u \frac{\partial}{\partial\theta} \sqrt{\alpha(s,\theta_1)}| \leq \Gamma \cdot \inf_{s\in[t_3,t_4]} |{}^t u \frac{\partial}{\partial\theta} \sqrt{\alpha(s,\theta_1)}|$

Remarques :

a) La condition B.ii) pourrait être remplacée comme dans [14], [15] par une plus générale consistant à supposer que $\alpha^{\overline{q}}$ est dérivable m fois et que toutes les dérivées sont de puissances $q^{\text{ièmes}}$ intégrables avec q entier pair satisfaisant $m.q > p$; cela alourdirait seulement les notations.

b) L'hypothèse C est l'habituel contrôle à l'infini avec l'affinité de Hellinger puisque avec des notations évidentes :

$$\rho(\theta_1+\theta, \theta_1 ; [t_1,t_2]) = e^{-\frac{1}{2}\int_{t_1}^{t_2} (\sqrt{\alpha(s,\theta_1+\theta)} - \sqrt{\alpha(s,\theta_1)})^2 ds}$$

c) Il est facile de voir que l'hypothèse D implique les conditions (4.3) et (4.4) page 191 de IBRAGIMOV-KHASMINSKII [23].

d) Nous verrons dans la suite que dans une certaine mesure (contribution à l'information, loi limite, ...), le gradient $\frac{\partial}{\partial\theta}\sqrt{\alpha(s,\theta_1)}$ joue le même rôle que le régresseur $f(.)$ dans [10] et il n'est pas étonnant de retrouver la condition F.

II.2. Convergence du processus de vraisemblance dans $\mathcal{C}(\mathbb{R}^p \times [0,1])$

Nous étudions la suite de processus de comptage N_n d'intensité : $n\ \alpha(.,\theta)$; le processus de vraisemblance normalisé (au point θ_o : vraie valeur du paramètre) est défini par le rapport des vraisemblances correspondant aux valeurs : $\theta_o + \frac{\theta}{\sqrt{n}}$ et θ_o du paramètre et $[0,t]$ comme période d'observation :

$$Z_n(\theta,t) = \exp \frac{1}{q} \left\{ \int_0^t \mathrm{Log} \frac{\alpha\,(s,\theta_o + \frac{\theta}{\sqrt{n}})}{\alpha\,(s,\theta_o)}\, dN_n\ (s) - \int_0^t n\,[\alpha\ (s,\theta_o + \frac{\theta}{\sqrt{n}}) - \right.$$

$$\left. - \alpha\ (s,\theta_o)]\, ds \right\}$$

avec q supérieur à p défini par l'hypothèse B. Par extension évidente, nous noterons $Z_n(.,\,[t_1,t_2]) = \frac{Z_n(.,\,t_2)}{Z_n(.,\,t_1)}$. Sauf quand cela sera nécessaire, nous ne noterons pas l'indice θ_o mais nous pouvons déjà affirmer que les convergences montrées dans la suite seront uniformes dans un voisinage de θ_o à cause de l'uniformité des hypothèses.

THEOREME 2.1.

Sous les hypothèses de régularité A,B, la suite de processus $Z_n(.,.)$ converge étroitement, sous P_{θ_o} , dans $\mathcal{C}(\mathbb{R}^p \times [0,1])$ muni de la topologie de la convergence uniforme sur les compacts vers le processus $Z(.,.)$ défini par :

$$Z(\theta,t) = \exp \frac{1}{q} \left\{ {}^t\theta \int_0^t \frac{\frac{\partial \alpha}{\partial \theta}\ (s,\theta_o)}{\sqrt{\alpha\ (s,\theta_o)}}\, dW\ (s) - \frac{1}{2}\ {}^t\theta\ I(\theta_o\ ;\ [0,t])\theta \right\}$$

où W est un mouvement brownien unidimensionnel sur $[0,1]$.

La convergence des marginales de dimension finie est immédiate car les processus N_n sont à accroissements indépendants et les hypothèses A sur l'intensité α permettent d'appliquer le théorème central limite pour les tableaux triangulaires : N_n peut être considéré comme la somme de n processus de Poisson indépendants, identiquement distribués.

La suite Z_n est contigüe à une suite de processus continus (en t) car log Z_n l'est et nous allons montrer la tension de cette suite Z_n pour la norme uniforme sur $[0,1] \times K$ où K est un compact quelconque de $\mathbb{R}^p$, à l'aide du lemme (2.1) mais démontrons d'abord le lemme préliminaire suivant :

LEMME (2.0) :

L'hypothèse B implique que la suite de processus $Z_n(.; [t_1,t_2])$ satisfait uniformément en n la "condition de mesure" utilisée dans le critère de tension de CSENSOV (voir par exemple [14]), plus précisément : il existe une constante C telle que $\forall t_1 < t_2$, $\forall \theta_o$, $\forall \theta_1$, $\forall \theta_2$,

$$E_{\theta_o} |Z_n(\theta_2;[t_1,t_2]) - Z_n(\theta_1,[t_1,t_2])|^q \leq C. \|\theta_2-\theta_1\|^q.(1 \vee \|\theta_o\| \vee \|\theta_1\| \vee \|\theta_2\|)^a$$

Démonstration :

Si θ_1 et θ_2 sont des points différant seulement par leur j^e coordonnée ($j = 1,\ldots,p$), nous pouvons écrire :

$$Z_n(\theta_2;[t_1,t_2]) - Z_n(\theta_1;[t_1,t_2])$$

$$= \int_{\theta_1^{(j)}}^{\theta_2^{(j)}} \frac{\partial}{\partial \theta^{(j)}} \exp \frac{1}{q} \left\{ \int_{t_1}^{t_2} \operatorname{Log} \frac{\alpha(s,\theta_o + \frac{\theta}{\sqrt{n}})}{\alpha(s,\theta_o)} dN_n(s) - \int_{t_1}^{t_2} n \left[\alpha(s,\theta_o + \frac{\theta}{\sqrt{n}}) - \alpha(s,\theta_o)\right] ds \right\} d\theta^{(j)}$$

En intégrant, on obtient de façon classique les majorations :

$$E_{\theta_o} |Z_n(\theta_2,[t_1,t_2]) - Z_n(\theta_1,[t_1,t_2])|^q$$

$$\leq \frac{|\theta_2^{(j)}-\theta_1^{(j)}|^q}{(q\sqrt{n})^q} \cdot \sup_{\theta_1^{(j)} \leq \theta^{(j)} \leq \theta_2^{(j)}} E_{\theta_o+\frac{\theta}{\sqrt{n}}} \left| \int_{t_1}^{t_2} \frac{\frac{\partial \alpha}{\partial \theta^{(j)}}(s,\theta_o+\frac{\theta}{\sqrt{n}})}{\alpha(s,\theta_o+\frac{\theta}{\sqrt{n}})} [dN_n(s) - n\ \alpha(s,\theta_o+\frac{\theta}{\sqrt{n}})\ ds] \right|^q$$

Le processus $N_n(.)$ pouvant être considéré comme la somme de n processus N_1 indépendants et identiquement distribués, le lemme (3.1) page 186 de [23] fournit la majoration :

$$C_q \cdot \|\theta_2-\theta_1\|^q \cdot \sup_{\theta \in [\theta_o+\frac{\theta_1}{\sqrt{n}},\ \theta_o+\frac{\theta_2}{\sqrt{n}}]} E_\theta \left| \int_{t_1}^{t_2} \frac{\frac{\partial \alpha}{\partial \theta^{(j)}}(s,\theta)}{\alpha(s,\theta)} \cdot [dN_1(s) - \alpha(s,\theta)ds] \right|^q$$

Le calcul des moments centrés pour un processus de Poisson implique (voir par exemple [31])

$$E_\theta \left| \int_{t_1}^{t_2} h(s)\ [dN_1(s)-\alpha(s,\theta)ds] \right|^q \leq C_1 \left(\int_{t_1}^{t_2} h(s)^2\ \alpha(s,\theta)ds \right)^{\frac{q}{2}} + C_2 \left(\int_{t_1}^{t_2} h(s)^q\ \alpha(s,\theta)ds \right)$$

et l'hypothèse B permet d'en déduire immédiatement le résultat cherché.

LEMME (2.1) :

Pour tout compact K de $\mathbb{R}^p$, et pour tout ε positif, il existe $\delta(\varepsilon,K)$ positif et un entier $n(\varepsilon,K)$ tels que $\forall \delta \leq \delta(\varepsilon,K)$, $\forall n \geq n(\varepsilon,K)$,

$$P\left\{ \sup_{\substack{t_1,t_2 \in [0,1] \\ |t_2-t_1| \leq \delta}} \sup_{\theta \in K} |Z_n(\theta,t_2)-Z_n(\theta,t_1)| \geq \varepsilon \right\} \leq \varepsilon \qquad (2.1)$$

$$P\left\{ \sup_{t \in [0,1]} \sup_{\substack{\theta_1,\theta_2 \in K \\ \|\theta_2-\theta_1\| \leq \delta}} |Z_n(\theta_2,t)-Z_n(\theta_1,t)| \geq \varepsilon \right\} \leq \varepsilon \qquad (2.2)$$

Démonstration de (2.1) :

Nous suivons la démonstration faite dans [14] pour le cas de variables indépendantes, identiquement distribuées en l'adaptant ici à un cas non stationnaire.

Pour un réel ε positif et un compact K de diamètre égal à D, nous découpons $[0,1]$ en ℓ intervalles égaux : ℓ sera fixé ultérieurement en fonction de ε et K.

$$P \Big\{ \sup_{\substack{t_1,t_2 \\ |t_2-t_1|\leq \frac{1}{\ell}}} \sup_{\theta\in K} |Z_n(\theta,t_2)-Z_n(\theta,t_1)| > \varepsilon \Big\}$$

$$\leq P\Big\{ \sup_{j=1,\ldots,\ell} \sup_{\frac{j-1}{\ell}\leq t\leq \frac{j}{\ell}} \sup_{\theta\in K} |Z_n(\theta,t)-Z_n(\theta,\tfrac{j-1}{\ell})| > \frac{\varepsilon}{3} \Big\}$$

$$\leq P\Big\{ \sup_{j=1,\ldots,\ell} \sup_{\theta\in K} Z_n(\theta,\tfrac{j-1}{\ell})>A\Big\} + \sum_{j=1}^{\ell} P\Big\{ \sup_{\frac{j-1}{\ell}\leq t\leq \frac{j}{\ell}} \sup_{\theta\in K} |Z_n(\theta;[\tfrac{j-1}{\ell},t])-1| > \frac{\varepsilon}{3A}\Big\}$$

où $A(\varepsilon,D)$ sera fixé ultérieurement.

$$\leq P\Big\{ \sup_{j=0,\ldots,\ell-1} \sup_{\theta\in K} \operatorname{Log} Z_n(\theta,\tfrac{j}{\ell}) > \operatorname{Log} A\Big\}$$

$$+ \sum_{j=0}^{\ell-1} P\Big\{ \sup_{\frac{j}{\ell}\leq t\leq \frac{j+1}{\ell}} \sup_{\theta\in K} |\operatorname{Log} Z_n(\theta,[\tfrac{j}{\ell},t])> \frac{\varepsilon}{6A}\Big\} = \pi_1+\pi_2 \ .$$

Chacun des deux termes est majoré à l'aide du lemme classique suivant que nous écrivons en temps continu :

LEMME (2.2): Soit $Y(t,\theta)$ un processus à accroissements indépendants en t sur $[0,1]$ et continu en θ sur un compact K de $\mathbb{R}^p$ à valeurs dans $\mathbb{R}^r$. Alors, $\forall A > 0$

$$P\Big\{ \sup_{t\in[0,1]} \sup_{\theta\in K} \|Y(t,\theta)\| > A\Big\}$$

$$\leq P\Big\{ \sup_{\theta\in K} \|Y(1,\theta)\| > \frac{A}{2}\Big\} \cdot \Big[1- \sup_{t\in[0,1]} P\Big\{\sup_{\theta\in K} \|Y(1,\theta)-Y(t,\theta)\| > \frac{A}{2}\Big\}\Big]^{-1}$$

Démonstration : Elle se fait de manière analogue à celle du cas où le paramètre t est discret,; il suffit de décomposer l'événement du premier membre suivant

les valeurs du temps d'arrêt $\tau_A = \inf \{t, \sup_{\theta\in K} \|Y(t,\theta)\| > A\}$.

On a:

$$\pi_1 \leq P \{\sup_{\theta\in K} \text{Log } Z_n(\theta,1)| \geq \frac{\text{Log } A}{2}\} \times$$

$$[1- \sup_{j=0,\ldots,\ell-1} P \{\sup_{\theta\in K} |\text{Log } Z_n(\theta,[\tfrac{j}{\ell}, 1])| \geq \frac{\text{Log } A}{2}\}]^{-1}$$

Pour s et t fixés $(s < t)$ la convergence de la suite de processus en $\theta: Z_n(.;[s,t])$ est uniforme sur tout compact car nous l'avons déjà dit, les marginales finies convergent vers celles de $Z(.,[s,t])$ et la tension dans $\mathscr{C}(K)$ est vérifiée grâce à l'hypothèse B assurant que la suite $Z_n(.,[s,t])$ satisfait la "condition de mesure" du critère de CSENSOV (voir [14]) par le lemme (2.0). Par conséquent :

$$P \{\sup_{\theta\in K} \text{Log } Z_n(\theta,1)| \geq \frac{\text{Log } A}{2}\} \leq 2P \{\sup_{\theta\in K} |\text{Log } Z(\theta,1)| \geq \frac{\text{Log } A}{2}\} \leq \frac{\varepsilon}{4}$$

pour $A \geq A_0(\varepsilon,K)$ et $n \geq n_0(\varepsilon,K,A_0)$.

De même, pour $A \geq A_1(\ell,K)$ et $n \geq n_1(\ell,K,A_1)$,

$$\sup_{j=0,\ldots,\ell-1} P \{\sup_{\theta\in K} |\text{Log } Z_n(\theta,[\tfrac{j}{\ell},1])| > \frac{\text{Log } A}{2}\} \leq \frac{1}{2} .$$

On majore aussi le second terme π_2 à l'aide du lemme (2.2) :

$$\pi_2 = \sum_{j=0}^{\ell-1} \{ \sup_{\frac{j}{\ell}\leq t\leq \frac{j+1}{\ell}} \sup_{\theta\in K} |\text{Log } Z_n(\theta,[\tfrac{j}{\ell},t])| > \frac{\varepsilon}{6A} \}$$

$$\leq \sum_{j=0}^{\ell-1} P\{\sup_{\theta\in K} |\text{Log } Z_n(\theta,[\tfrac{j}{\ell},\tfrac{j+1}{\ell}])| > \frac{\varepsilon}{12A}\} \times [1- \sup_{\frac{j}{\ell}\leq t\leq \frac{j+1}{\ell}} P \{\sup_{\theta\in K} \text{Log } Z_n(\theta,[t,\tfrac{j+1}{\ell}])| > \frac{\varepsilon}{12A}]^{-1}$$

La convergence à ℓ fixé de la suite de processus $Z_n(.,[\frac{j}{\ell},\frac{j+1}{\ell}])$ permet d'écrire :

$$\sum_{j=0}^{\ell-1} P\{ \sup_{\theta,\|\theta\| \leq D} |\text{Log } Z_n(\theta,[\tfrac{j}{\ell}, \tfrac{j+1}{\ell}])| > \tfrac{\varepsilon}{12A}\}$$

$$\leq 2 \sum_{j=0}^{\ell-1} P\{ \sup_{\theta,\|\theta\| \leq D} |\text{Log } Z(\theta,[\tfrac{j}{\ell}, \tfrac{j+1}{\ell}])| > \tfrac{\varepsilon}{12A}\}$$

$$= 2 \sum_{j=0}^{\ell-1} P\{ \sup_{\theta,\|\theta\| \leq D} |{}^t\theta\xi_j - \tfrac{1}{2}\, {}^t\theta\, V_j\theta| > \tfrac{\varepsilon}{12A}\}$$

où ξ_j est un vecteur gaussien centré de variance $V_j = I(\theta_o;[\tfrac{j}{\ell}, \tfrac{j+1}{\ell}])$

Pour ℓ sufisamment grand, le suprémum est atteint au bord de la sphère et $\sup_j \frac{1}{2} D^2 \|V_j\| \leq \frac{\varepsilon}{24A}$, ce qui fournit le majorant plus simple :

$$2 \sum_{j=0}^{\ell-1} P\{\|\xi_j\| \geq \tfrac{\varepsilon}{24\,AD}\}$$

$$\leq 2 \sum_{j=0}^{\ell-1} (\tfrac{24\,AD}{\varepsilon})^4 \, E\|\xi_j\|^4 \leq \tfrac{\varepsilon}{4} \quad \text{pour} \quad \ell \geq \ell_o(A,\varepsilon,D)$$

Pour achever la majoration de π_2 et donc la démonstration de (2.1) , il suffit de prouver que

$$\sup_{\substack{s,t\in[0,1] \\ |t-s| \leq \frac{1}{\ell}}} P\{ \sup_{\theta,\|\theta\| \leq D} |\text{Log } Z_n(\theta,[s,t])| > \tfrac{\varepsilon}{12A}\} \leq \tfrac{1}{2}$$

ce qui est vrai uniformément en n, pour ℓ assez grand grâce au lemme (2.0).

Démonstration de l'assertion (2.2) :

Pour ε positif et un compact K pris égal à $[0,D]^p$ pour simplifier, nous découpons ce compact en pavés de côtés δ , d'extrémités $\underline{\theta}_j$ et $\overline{\theta}_j$, $j \in J$; nous avons alors la suite de majorations :

$$P\{\sup_{t\in[0,1]}\ \sup_{\substack{\theta_1,\theta_2\in[0,D]^p\\ \|\theta_2-\theta_1\|\le\delta}} |Z_n(\theta_2,t) - Z_n(\theta_1,t)| \ge \varepsilon\}$$

$$\le P\{\sup_{t\in[0,1]}\ \sup_{j\in J}\ \sup_{\theta\in[\underline{\theta}_j,\overline{\theta}_j]} |Z_n(\theta,t)-Z_n(\underline{\theta}_j,t)| \ge \frac{\varepsilon}{3}\}$$

$$\le P\{\sup_{t\in[0,1]}\ \sup_{j\in J} Z_n(\underline{\theta}_j;t) \ge A\}$$

$$+ P\{\sup_{t\in[0,1]}\ \sup_{j\in J}\ \sup_{\theta\in[\underline{\theta}_j,\overline{\theta}_j]} \left|\frac{Z_n(\theta,t)}{Z_n(\underline{\theta}_j,t)} - 1\right| \ge \frac{\varepsilon}{3A}\}$$

où $A(\varepsilon,D)$ sera fixé ultérieurement.

$$\le P\{\sup_{t\in[0,1]}\ \sup_{j\in J} \text{Log } Z_n(\underline{\theta}_j,t) \ge \text{Log } A\}$$

$$+ P\{\sup_{t\in[0,1]}\ \sup_{j\in J}\ \sup_{\theta\in[\underline{\theta}_j,\overline{\theta}_j]} \left|\text{Log}\frac{Z_n(\theta,t)}{Z_n(\underline{\theta}_j,t)}\right| > \frac{\varepsilon}{6A}\} = \pi_1' + \pi_2' .$$

Nous utilisons le lemme (2.2) pour majorer π_1' d'abord :

$$\pi_1' \le P\{\sup_{j\in J} |\text{Log } Z_n(\underline{\theta}_j,1)| \ge \frac{\text{Log } A}{2}\}\times[1-\sup_{t\in[0,1]} P\{\sup_{j\in J} |\text{Log } Z_n(\underline{\theta}_j,[t,1])| \ge \frac{\text{Log } A}{2}\}]^{-1}$$

Le numérateur est encore majoré par :

$$P\{\sup_{\theta\in[0,D]^p} |\text{Log } Z_n(\theta,1)| \ge \frac{\text{Log}A}{2}\} \le \frac{\varepsilon}{4}$$

car c'est le même numérateur que π_1. Le dénominateur de π_1' sera minoré par $\frac{1}{2}$ si on montre que

$$\sup_{t\in[0,1]} P\{\sup_{\theta\in[0,D]^p} \left|\text{Log}\frac{Z_n(\theta,1)}{Z_n(\theta,t)}\right| \ge \frac{\text{Log } A}{2}\} \le \frac{1}{2} .$$

Cette dernière inégalité est satisfaite uniformément en n, pour A suffisamment grand en utilisant le lemme (2.0).

$$\pi_2' \leq P\{\sup_{j\in J} \sup_{\theta\in[\underline{\theta}_j,\overline{\theta}_j]} \left|\mathrm{Log}\frac{Z_n(\theta,1)}{Z_n(\underline{\theta}_j,1)}\right| > \frac{\varepsilon}{12A}\} \times$$

$$\left[1-\sup_{t\in[0,1]} P\{\sup_{j\in J} \sup_{\theta\in[\underline{\theta}_j,\overline{\theta}_j]} \left|\mathrm{Log}\frac{Z_n(\theta;[t,1])}{Z_n(\overline{\theta}_j,[t,1])}\right| > \frac{\varepsilon}{12A}\right]^{-1}$$

$$\leq P\{\sup_{j\in J} \sup_{\theta\in[\underline{\theta}_j,\overline{\theta}_j]} \left|\frac{Z_n(\theta,1)}{Z_n(\underline{\theta}_j,1)} - 1\right| > \frac{\varepsilon}{12A}\} \times$$

$$\left[1-\sup_{t\in[0,1]} P\{\sup_{j\in J} \sup_{\theta\in[\underline{\theta}_j,\overline{\theta}_j]} \left|\frac{Z_n(\theta;[t,1])}{Z_n(\overline{\theta}_j,[t,1])} - 1\right| > \frac{\varepsilon}{12A}\}\right]^{-1}$$

Le numérateur de ce dernier majorant est plus petit que :

$$P\{\inf_{j\in J} Z_n(\underline{\theta}_j,1) \leq \frac{1}{\sqrt{A}}\} + P\{\sup_{j\in J} \sup_{\theta\in[\underline{\theta}_j,\overline{\theta}_j]} |Z_n(\theta,1)-Z_n(\underline{\theta}_j,1)| \geq \frac{\varepsilon}{12A^{3/2}}\}$$

Comme précédemment, la convergence de la suite $Z_n(.;1)$ vers $Z(.,1)$ dans $\mathscr{C}([0,D]^p)$ suffit à rendre cette somme inférieure à $\frac{\varepsilon}{8}$ pour n assez grand. Il suffit pour achever la preuve de (2.2) de vérifier que

$$\sup_{t\in[0,1]} P\{\inf_{j\in J} Z_n(\underline{\theta}_j;[t,1]) \leq \frac{1}{\sqrt{A}}\} + P\{\sup_{j\in J} \sup_{\theta\in[\underline{\theta}_j,\overline{\theta}_j]} |Z_n(\theta,[t,1])-Z_n(\underline{\theta}_j,[t,1])| > \frac{\varepsilon}{12A^{3/2}}\} \leq \frac{3}{4}$$

On a déjà montré que :

$$\sup_{t\in[0,1]} P\{\sup_{\theta\in[0,D]^p} |\mathrm{Log}\, Z_n(\theta,[t,1]|) > \frac{\mathrm{Log}\, A}{2}\} \leq \frac{1}{2}$$

Il reste à prouver que pour n assez grand

$$\sup_{t\in[0,1]} P\{\sup_{j\in J} \sup_{\theta\in[\underline{\theta}_j,\overline{\theta}_j]} |Z_n(\theta,[t,1])-Z_n(\underline{\theta}_j,[t,1])| > \frac{\varepsilon}{12A^{3/2}}\} \leq \frac{1}{4}$$

or ceci est vrai uniformément en n, pour δ suffisamment petit grâce au lemme (2.0).

II.3. Convergence du processus de vraisemblance hors des compacts.

Le contrôle à l'infini (en θ) de la suite $Z_n(.,.)$ consiste en la convergence dans un espace $\mathscr{C}_\varphi(\mathbb{R}^p x[0,1])$ remplaçant l'espace $\mathscr{C}_0$; la convergence vers 0 hors des compacts en θ ne peut être uniforme en t à cause du voisinage de 0 : $Z_n(.,o) \equiv 1$.

Nous aurons besoin de l'hypothèse supplémentaire :

$\underline{H}$: il existe $r > 0$, $t_o > 0$ et un entier $n_o > 0$ tels que pour tout θ_1 appartenant à un voisinage de θ_o ,

$$E_{\theta_1} \sup_{\theta \in \mathbb{R}^p} Z_{n_1}(\theta,t_o)^r < \infty \qquad \text{et} \qquad E_{\theta_1} \sup_{\theta \in \mathbb{R}^p} Z_{n_o}(\theta;[1-t_o,1])^r < \infty \ .$$

PROPOSITION (2.2) :

Sous les hypothèses A,B,C,F,H et pour une fonction φ définie sur $[0,1]$ satisfaisant la condition :

$\underline{E}$: $\exists c > 0$ et $\exists \beta > 0$ tels que $\varphi(t) \leq c.t^\beta$, $\forall t \in [0,1]$

On a : $\forall \varepsilon > 0$, $\exists M > 0$ et $\exists n_o(\varepsilon)$ tel que : $\forall n \geq n_o$,

$$P_{\theta_1}\{\sup_{t\in]0,1]} \sup_{\theta,\|\theta\|\varphi(t)\geq M} Z_n(\theta,t) \geq \varepsilon\} \leq \varepsilon \ .$$

La démonstration va être faite en une succession de lemmes qui prouveront que cette majoration est uniforme en θ_1 appartenant à un voisinage de la "vraie valeur" θ_o .

LEMME (2.3) :

Sous les hypothèses A,B,C,F, la suite de processus Z_n satisfait : $\forall \nu > 0$, $\exists B > 0$ et $\exists n_o(\nu)$ tels que : $\forall t_1,t_2 \in [0,1]$, $t_1 \leq t_2$, $\forall \theta_1$, $\|\theta_1-\theta_o\| < \delta'$, $\forall n \geq n_o$

$$E_{\theta_1} Z_n(\theta;[t_1,t_2])^{\frac{q}{2}} \leq \frac{B}{\{{}^t\theta[I(\theta_1,[t_1,t_2])]\,\theta\}^{\nu}}$$

Démonstration du lemme (2.3) :

Le calcul direct fournit :

$$E_{\theta_1} Z_n(\theta;[t_1,t_2])^{\frac{q}{2}} = e^{-\frac{n}{2}\int_{t_1}^{t_2}\left(\sqrt{\alpha(s,\theta_1+\frac{\theta}{\sqrt{n}})}-\sqrt{\alpha(s,\theta_1)}\right)^2 ds}$$

D'après l'hypothèse C, il existe une constante $\delta > 0$ telle que $\forall\theta \in \mathbb{R}^p$, $\forall t_1,t_2 \in [0,1]$,

$$\{{}^t\theta\, I(\theta_1,[t_1,t_2])\theta\}^{\delta}\; e^{-\frac{1}{2}\int_{t_1}^{t_2}\left(\sqrt{\alpha(s,\theta_1+\theta)}-\sqrt{(s,\theta_1)}\right)^2 ds} \leq 1 \ .$$

On en déduit de façon classique ([22]) l'existence pour tout ν positif d'une constante U positive et d'un entier n_1 tels que : $\forall n \geq n_1, \forall\theta$, $\forall t_1$, $\forall t_2$, ${}^t\theta\, I(\theta_1,[t_1,t_2])\theta > U$,

$$E_{\theta_1} Z_n(\theta;[t_1,t_2])^{\frac{q}{2}} \leq \frac{1}{\{{}^t\theta\, I(\theta_1;[t_1,t_2])\theta\}^{\nu}}$$

Pour les valeurs de θ , t_1 et t_2 satisfaisant : $1 \leq {}^t\theta\, I(\theta_1,[t_1,t_2])\theta \leq U$, on utilise le comportement de $\sqrt{\alpha(.,.)}$ au voisinage d'un point, donné par les hypothèses A, B et F : U étant fixé, il existe $\delta"$ positif tel que :

$$E_{\theta_1} Z_n(\theta;[t_1,t_2])^{\frac{q}{2}} \leq e^{-\delta"\,{}^t\theta\, I(\theta_1,[t_1,t_2])\theta}$$

et par conséquent, l'existence d'une constante B telle que : $\forall\, u \leq U$,

$$e^{-\delta" u^2} \leq \frac{B}{u^{\nu}}$$

implique le résultat du lemme (2.3).

Remarque : Ce dernier lemme fournit une majoration identique à celle obtenue dans le lemme (4.3) de [15] pour le cas d'une régression : le vecteur $\frac{\partial}{\partial\theta}\sqrt{\alpha(.,\theta_o)}$ joue le rôle du régresseur $f(.)$ et l'information de Fisher $I(\theta_o;.)$ remplace la matrice $F(.)$.

La conjonction des lemmes (2.0) et (2.3) permet de déduire en utilisant l'inégalité de Holder le :

LEMME (2.4) :

Sous les hypothèses A,B,C, $\forall\nu' > 0$, $\exists B' > 0$, $\exists q' \in]p,q[$, $\exists a' > 0$ et $\exists n'_o(\nu')$ tels que : $\forall n \geq n'_o$, $\forall\theta_1$, $\|\theta_1-\theta_o\| \leq \delta'$, $\forall\theta_2$, $\forall\theta_3 \in \mathbb{R}^p$, $\forall t_1$, $\forall t_2$

$$E_{\theta_1} |Z_n(\theta_3,[t_1,t_2]) - Z_n(\theta_2,[t_1,t_2])|^{q'} \leq \frac{B''_o \|\theta_3-\theta_2\|^{q'}}{\underset{\theta\in[\theta_2,\theta_3]}{\text{Inf}} \{{}^t\theta\, I(\theta_1,[t_1,t_2])\theta\}^{\nu'}}$$

Ce résultat apporte une majoration uniforme en n de la suite de processus $Z_n(.,.[t_1,t_2])$ sur des cubes de coté 1 par le lemme de CSENSOV (voir le lemme (1.3) de [14]) : il existe B'' tel que $\forall A > 0$, $\forall n \geq n'_o$

$$P_{\theta_1} \{\sup_{\theta,\theta\in[\theta_2,\theta_3]} Z_n(\theta,[t_1,t_2]) > A\} \leq \frac{B''}{A^{q'}. \underset{\theta\in[\theta_2,\theta_3]}{\text{Inf}} \{{}^t\theta\, I(\theta_1,[t_1,t_2])\theta\}^{\nu'}}$$

Par sommation sur les cubes élémentaires, on montre le contrôle à l'infini de la suite de processus $Z_n(.;[t_1,t_2])$ pour t_1 et t_2 fixés :

LEMME (2.5) :

Il existe une constante B''' telle que : $\forall\ell$, $\forall A > 0$, $\forall t_1$, $\forall t_2$, $\forall\theta_1$ $\|\theta_1-\theta_o\| < \delta'$

$$P_{\theta_1} \{\sup_{\theta,\,{}^t\theta\, I(\theta_1,[t_1,t_2])\theta>\ell} Z_n(\theta;[t_1,t_2]) \geq A\} \leq B''' .A^{-q'}.\ell^{\frac{p}{2}-\nu'} \text{ dét } (I(\theta_1;[t_1,t_2]))^{-\frac{1}{2}}$$

Ce résultat est analogue à celui du lemme (4.5) de [15] ; la démonstration de la proposition (2.2) s'achève de façon identique à [15] : il faut découper l'événement considéré en deux parties : l'une

$$\{\sup_t \sup_{\theta,\ \frac{M}{\varphi(t)} < \|\theta\| \leq \sqrt{n}\,\delta} Z_n(\theta,t) \geq \varepsilon\}$$

où nous pouvons faire un développement de Taylor au 2ème ordre de $\log Z_n$ et utiliser l'hypothèse D ; l'autre

$$\{\sup_t \sup_{\theta,\|\theta\| > \frac{M}{\varphi(t)} \vee \sqrt{n}\,\delta} Z_n(\theta,t) \geq \varepsilon\}$$

où nous pouvons appliquer le lemme (2.5) ci-dessus.

La proposition suivante est une conséquence du résultat précédent et sera utile pour les tests.

PROPOSITION (2.3) :

Sous les hypothèses A,B,C,D,E,F,H la suite de processus

$$S_n(t) = \varphi(t) \sup_{\theta\in\mathbb{R}^p} \operatorname{Log} Z_n(\theta,t)$$

converge étroitement sous P_{θ_o} vers le processus

$$S(t) = \varphi(t) \sup_{\theta\in\mathbb{R}^p} \operatorname{Log} Z(\theta,t)$$

Démonstration :

Par la proposition précédente, pour tout ε positif nous avons la convergence dans $\mathscr{C}_0(\mathbb{R}^p \times [\varepsilon,1])$ de la suite $Z_n(.,.)$ vers Z en prenant $\varphi(s) = 1_{[\varepsilon,1]}(s)$ et on obtient ainsi la convergence dans $\mathscr{C}[\varepsilon,1]$ de la suite $S_n(.)$ vers $S(.)$.

Il suffit alors de montrer la tension en 0 de la suite $S_n(.)$, c'est-à-dire : $\forall\varepsilon' > 0$, $\exists\varepsilon > 0$ et $\exists n_o$ tels que $\forall n \geq n_o$,

$$P\{\sup_{t\in[0,\varepsilon]} \varphi(t) \operatorname{Log} \sup_{\theta\in\mathbb{R}^p} Z_n(\theta,t) \geq \varepsilon'\} \leq \varepsilon'$$

Pour tout M positif, la convergence uniforme sur les compacts de $Z_n(.,.)$ entraîne l'existence de ε_1 vérifiant :

$$P\{\sup_{t\in[0,\varepsilon_1]} \sup_{\theta,\|\theta\|\leq M} Z_n(\theta,t) \geq e^{\frac{\varepsilon'}{\|\varphi\|}}\} \leq \frac{\varepsilon'}{2}$$

Il reste donc à trouver ε suffisamment petit et M suffisamment grand qui satisferont pour n assez grand :

$$P\{\sup_{t\in[0,\varepsilon]} \sup_{\theta,\|\theta\|>M} \varphi(t) \operatorname{Log} Z_n(\theta,t) \geq \varepsilon'\} \leq \frac{\varepsilon'}{2} .$$

La démarche est la même que dans [15] et consiste encore à scinder cet événement en deux : le premier

$$\{\sup_{t\in[0,\varepsilon]} \sup_{\theta,M<\|\theta\|\leq\sqrt{n}\delta_0} \varphi(t) \operatorname{Log} Z_n(\theta,t) \geq \varepsilon'\}$$

où un développement de Taylor au 2ème ordre conduit au résultat et le second :

$$\{\sup_{t\in[0,\varepsilon]} \sup_{\theta,\|\theta\|>M\vee\sqrt{n}\delta_0} \varphi(t) \operatorname{Log} Z_n(\theta,t) \geq \varepsilon'\}$$

où on applique le lemme (2.5).

II.4. Test du rapport de vraisemblance modifié et approximation du niveau.

Comme dans le cas d'un n-échantillon ou d'un modèle de régression, la proposition précédente met en évidence la nécessité de renormaliser la statistique du test du rapport de vraisemblance en pénalisant les bords de l'intervalle d'observation $[0,1]$ par une fonction $\varphi(.)$. Le rapport de vraisemblance correspondant à la j^e coordonnée d'un processus $N_n=(N_n^{(1)},\ldots N_n^{(2)})$ s'écrit en un point θ_o de Θ_o :

$$\Lambda_n^{(j)}(\theta^*,\theta^{**},t;\theta) = \left[\frac{Z_n^{(j)}(\theta^*;[0,t]).\ Z_n^{(j)}(\theta^{**};[t,1])}{Z_n^{(j)}(\theta;1)} \right]^q$$

PROPOSITION 2.4. :

Pour une suite $N_n = (N_n^{(1)},\ldots, N_n^{(r)})$ de processus dont les coordonnées ont des intensités $\alpha^{(j)}$ satisfaisant les hypothèses A, B, C, D, F et H pour toute fonction φ satisfaisant la condition E, la suite de statistiques

$$\sup_{t\in]0,1[} \sup_{\theta^*\in\mathbb{R}^p} \sup_{\theta^{**}\in\mathbb{R}^p} \inf_{\theta\in\mathbb{R}^p} \varphi(t) \sum_{j=1}^{r} \text{Log}\ \Lambda_n^{(j)}(\theta^*,\theta^{**},t;\theta)$$

converge sous $H_{0,n}(\theta_o)$ vers

$$\frac{1}{2} \sup_{t\in]0,1[} \varphi(t) \Big\{ {}^t\Big[\int_0^t \Delta(s,\theta_o)\, dW(s)\Big] \Big[\int_0^t \Delta(s,\theta_o)\, {}^t\Delta(s,\theta_o)\, ds\Big]^{-1} \Big[\int_0^t \Delta(s,\theta_o)\, dW(s)\Big]$$

$$+ {}^t\Big[\int_t^1 \Delta(s,\theta_o)\, dW(s)\Big] \Big[\int_t^1 \Delta(s,\theta_o)\, {}^t\Delta(s,\theta_o)\, ds\Big]^{-1} \Big[\int_t^1 \Delta(s,\theta_o)\, dW(s)\Big]$$

$$- {}^t\Big[\int_0^1 \Delta(s,\theta_o)\, dW(s)\Big] \Big[\int_0^1 \Delta(s,\theta_o)\, {}^t\Delta(s,\theta_o)\, ds\Big]^{-1} \Big[\int_0^1 \Delta(s,\theta_o)\, dW(s)\Big]$$

où $\Delta(s,\theta_o)$ est la matrice (p,r) des dérivées partielles :

$$\Big(\frac{\partial}{\partial\theta}\sqrt{\alpha^{(1)}(s,\theta_o)},\ldots,\ \frac{\partial}{\partial\theta}\sqrt{\alpha^{(r)}(s,\theta_o)}\Big)$$

et W est un mouvement brownien r-dimensionnel sur $[0,1]$.

Démonstration : Elle repose sur la décomposition à t fixé de $\text{Log}\ \Lambda_n^{(j)}$ en trois termes tels que chacun multiplé par $\varphi(t)$ converge dans $\mathscr{C}[0,1]$ d'après la proposition (2.3) ; le passage au cas vectoriel $(r > 1)$ étant immédiat par l'indépendance des coordonnées du processus $(N_n^{(1)},\ldots, N_n^{(r)})$. Le reste de la démonstration n'est que le calcul des supréma-infima en θ^*, θ^{**}, θ sur le processus limite.

Remarque : La proposition précédente est bien souvent sans intérêt pour la pratique : approximation du niveau ou recherche d'un seuil pour le test du rapport de vraisemblance (modifié par φ) car la loi limite n'est pas libre sous H_o. Cependant, dans le cas où l'intensité se factorise de façon particulière, on peut énoncer le corollaire suivant :

COROLLAIRE (2.5) :

Soit une suite $N_n = (N_n^{(1)},\ldots, N_n^{(r)})$ de processus de comptage dont les intensités $\alpha_j(.,\theta)$, $j=1,\ldots,r$ sont de la forme :

$$\sqrt{\alpha_j(s,\theta)} = {}^tG(\theta)\, f_j(s) + b_j(s)$$

avec $G(\theta)$: vecteur de même dimension p que θ et de matrice jacobienne inversible.

$b_j(s)$: fonction de carré intégrable sur $[0,1]$

$f_j(s)$: fonction vectorielle de dimension p

et satisfont les conditions A,B,C,D,F,H précédentes.

Pour une fonction φ satisfaisant la condition E,

$$\sup_{t\in]0,1[} \varphi(t) \sup_{\theta^*\in\mathbb{R}^p} \sup_{\theta^{**}\in\mathbb{R}^p} \inf_{\theta\in\mathbb{R}^p} \sum_{j=1}^{r} \mathrm{Log}\, \Lambda_n^{(j)}(\theta^*,\theta^{**},t;\theta)$$

converge sous $H_{0,n}(\theta_o)$ vers :

$$\frac{1}{2} \sup_{t\in]0,1[} \varphi(t) \{ {}^t\Big[\int_o^t f(s)\, dW(s)\Big]\Big[\int_o^t f(s)\, {}^tf(s)\, ds\Big]^{-1}\Big[\int_o^t f(s)\, dW(s)\Big]$$

$$+ {}^t\Big[\int_o^1 f(s)\, dW(s)\Big]\Big[\int_t^1 f(s)\, {}^tf(s)\, ds\Big]^{-1}\Big[\int_t^1 f(s)\, dW(s)\Big]$$

$$- {}^t\Big[\int_o^1 f(s)\, dW(s)\Big]\Big[\int_o^1 f(s)\, {}^tf(s)\, ds\Big]^{-1}\Big[\int_o^1 f(s)\, dW(s)\Big]\}$$

où $f(.) = (f_1(.),\ldots,f_r(.))$.

La démonstration consiste seulement à éliminer les matrices jacobiennes $G'(\theta_o)$ qui interviennent dans $\Delta(s,\theta_o)$, ce qui est possible car elles sont inversibles.

Remarques :

i) Dans le cas particulier du corollaire avec de plus $p=1$, il est inutile de considérer $N_n^{(1)},\ldots,N_n^{(r)}$ puisque $\bar{N}_n = \sum_{i=1}^{r} N_n^{(i)}$ est exhaustif pour les paramètres (sous H_0 et sous H_1) et nous sommes ramenés en fait à étudier le cas stationnaire par un changement de temps (voir I). Toutefois, il serait utile de conserver toutes les coordonnées si on envisageait comme alternative le changement de paramètre d'une seule coordonnée.

ii) Le résultat de ce corollaire est à rapprocher de celui obtenu dans [15] pour un modèle de translation-régression sur le temps : la matrice d'information est la même.

II.5. *Loi asymptotique des estimateurs du maximum de vraisemblance.*

Il s'agit en présence d'une rupture d'estimer les trois paramètres $(\theta^*,\theta^{**},\tau)$ la caractérisant. Dans le cas stationnaire, la loi asymptotique des estimateurs du maximum de vraisemblance a été établie lorsque n tend vers l'infini, les paramètres $\theta^*,\theta^{**},\tau$ restant fixes ; la loi asymptotique de $\hat{\tau}_n$ dépend alors trop fortement du rapport $\rho = \frac{\theta^{**}}{\theta^*}$ ou de leur différence : $d = \theta^{**}-\theta^*$: elle nécessite un nouveau calcul de loi pour toute valeur de ρ (ou de d). Un moyen d'obtenir une seule loi limite est de considérer une suite d'expériences où le paramètre τ reste fixe mais où la différence $\theta_n^{**} - \theta_n^* = d_n$ tend vers 0 lorsque n tend vers l'infini. Comme dans [15], on suppose aussi que l'amplitude de la rupture d'intensité n'est pas trop petite: $\sqrt{n}[\alpha(\tau,\theta_n^{**})-\alpha(\tau,\theta_n^*)]$ tend vers l'infini ce qui implique, en particulier, que $\sqrt{n}\, d_n$ tend vers l'infini, ceci afin d'éviter d'être sous une suite d'hypothèses contigues à H_0 : cas où l'instant de rupture ne peut être estimé de façon consistante. L'hypothèse $H_{1,n}$ $(\theta_n^*,\theta_n^{**};\tau)$ se paramétrise encore par $\theta_n^{**} -\theta_n^*$ et par $\tau\theta_n^* + (1-\tau)\,\theta_n^{**}$ qui peut être supposé constant et égal à θ_o :

$$\theta_n^* = \theta_o - (1-\tau)\, d_n$$

$$\theta_n^{**} = \theta_o + \tau \cdot d_n$$

La loi asymptotique des estimateurs de ces paramètres va être établie en montrant d'abord la convergence d'un nouveau processus de vraisemblance sous $H_{1,n}\ (\theta_n^*, \theta_n^{**}, \tau)$:

$$\pi_n(\theta^*,\theta^{**},T) = \frac{dP_{H_{1,n}}\left(\theta_n^* + \frac{\theta^*}{\sqrt{n}}\ ;\ \theta_n^{**} + \frac{\theta^{**}}{\sqrt{n}}\ ;\ \tau + \frac{T}{\nu_n}\right)}{dP_{H_{1,n}}\left(\theta_n^*, \theta_n^{**}, \tau\right)} (N_n)$$

défini sur $\mathbb{R}^p \times \mathbb{R}^p \times [-\nu_n\tau,\ \nu_n(1-\tau)]$ et prolongé par 0 en dehors de cet ensemble de $\mathbb{R}^{2p+1}$ avec $\nu_n = n\ \left[{}^t d_n\ \frac{\partial}{\partial\theta}\ 2\ \overline{\sqrt{\alpha(\tau,\theta_o)}}\right]^2$ tendant vers l'infini.

THEOREME (2.6) :

Pour une suite de processus N_n de processus dont l'intensité $n\ \alpha(.,.)$ satisfait les hypothèses A,B,C,D,F,H, la suite des processus de vraisemblance π_n converge étroitement sous $H_{1,n}\ (\theta_n^*, \theta_n^{**}, \tau)$ dans $\mathscr{C}_o\ (\mathbb{R}^{2p+1})$ vers le processus π défini par :

$$\pi\ (\theta^*,\theta^{**},T) = \exp\ \{{}^t\theta^*\ I(\theta_o;[0,\tau])^{\frac{1}{2}}\ \xi_1 + {}^t\theta^{**}\ I(\theta_o;[\tau,1])^{\frac{1}{2}}\ \xi_2 + W(T)$$

$$- \frac{1}{2} \cdot {}^t\theta^*\ I(\theta_o;[0,\tau])\theta^* - \frac{1}{2}\ {}^t\theta^{**}\ I(\theta_o;[\tau,1])\ \theta^{**} - \frac{|T|}{2}\}$$

où ξ_1 et ξ_2 sont des gaussiennes indépendantes $\mathscr{N}(0;\ I_{p,p})$ et indépendantes de $W(.)$: mouvement brownien sur $\mathbb{R}$.

COROLLAIRE (2.7) :

Sous les mêmes hypothèses, les estimateurs du maximum de vraisemblance de $\theta_n^*, \theta_n^{**}, \tau$ sont asymptotiquement indépendants et

$$\sqrt{n}\ (\hat{\theta}_n^* - \theta_n^*) \xrightarrow[n\to\infty]{} \mathcal{N}(0,\ I(\theta_o;[0,\tau])^{-1})$$

$$\sqrt{n}\ (\hat{\theta}_n^{**} - \theta_n^{**}) \xrightarrow[n\to\infty]{} \mathcal{N}(0,\ I(\theta_o;\ [\tau,1])^{-1})$$

$n\ [{}^t d_n \frac{\partial}{\partial\theta}\ 2\ \sqrt{\alpha(\tau,\theta_o)}]^2\ (\hat{\tau}_n - \tau) \xrightarrow[n\to\infty]{} \zeta$ où ζ désigne la loi de l'instant de suprémum sur $\mathbb{R}$ de $W(T) - \frac{|T|}{2}$.

Démonstration : La déduction du corollaire à partir du théorème est immédiate et est en tous points identique à celle faite dans [15] pour le modèle de régression. Montrons le théorème.

Les paramètres avant et après rupture jouant le même rôle, nous nous contenterons de montrer la convergence de la suite de processus π_n pour $T \geq 0$. En reprenant les notations Z_n du processus de vraisemblance défini cette fois aux points θ_n^* et θ_n^{**} au lieu d'un θ_o fixé, π_n se décompose en trois processus indépendants séparant presque les paramètres :

$$\pi_n(\theta^*,\theta^{**},T)^{\frac{1}{q}} = Z_{n,\theta_n^*}\ (\theta^*;[0,\tau]\ ).\ Z_{n,\theta_n^{**}}\ (\theta^* - \sqrt{n}\ d_n;[\tau,\tau + \frac{T}{\nu_n}])$$

$$\times\ Z_{n,\theta_n^{**}}\ (\theta^{**};\ [\tau + \frac{T}{\nu_n},1])$$

La convergence uniforme de π_n sur les compacts de $\mathbb{R}^{2p+1}$ va être assurée par celle de chacun des trois termes (ou de leur logarithme).

Les hypothèses de départ impliquent que la convergence de $Z_{n,\theta_o}(.,.)$ est uniforme en θ_o appartenant à un compact donc en particulier pour des suites θ_n^* et θ_n^{**} convergeant vers θ_o. On a donc la convergence du 1er terme : $Z_{n,\theta_n^*}\ (\theta^*;[0,\tau]\ )$ dans $\mathcal{C}(\mathbb{R}^p)$ vers $e^{{}^t\theta^*\ I(\theta_o;[0,\tau])^{1/2}\ \xi_1 - \frac{1}{2}\ {}^t\theta^*\ I(\theta_o;[0,\tau])\theta^*}$. On a de même la convergence de

$Z_{n,\theta_n^{**}}(\theta^{**}; [\tau,1]$ dans $\mathcal{C}(\mathbb{R}^p)$ vers $e^{{}^t\theta^{**}I(\theta_o;[\tau,1])^{1/2}\xi_2 - \frac{1}{2}{}^t\theta^{**}I(\theta_o;[\tau,1])\theta^{**}}$

ce qui entrainera la convergence du 3ème terme si on montre le lemme suivant :

LEMME (2.6) : Sous les hypothèses du théorème (2.6), la suite de processus

$$Z_{n,\theta_n^{**}}(\theta^{**}, [\tau,\tau + \frac{T}{\nu_n}])$$

converge uniformément vers 1 sur les compacts de $\mathbb{R}^p \times \mathbb{R}^+$.

Démonstration du lemme :

Soit $\varepsilon > 0$ et $K \times [0,\bar{T}]$ un compact de $\mathbb{R}^{p+1}$. Pour le processus limite

$$Z_{\theta_o}(\theta^{**}, [\tau,t]) = \exp \frac{1}{q}\left\{{}^t\theta^{**}\int_\tau^t \frac{\partial}{\partial\theta} 2\sqrt{\alpha(s,\theta_o)}dW(s) - \frac{1}{2}{}^t\theta^{**}I(\theta_o;[\tau,t])\,\theta^{**}\right\}$$

il existe $\delta(K,\varepsilon)$ tel que :

$$P\left\{\sup_{\theta^{**}\in K}\ \sup_{\tau\le t\le \tau+\delta} |Z_{\theta_o}(\theta^{**}, [\tau,t])-1| > \frac{\varepsilon}{2}\right\} \le \frac{\varepsilon}{2}$$

La convergence de $Z_{n,\theta_n^{**}}(\theta^{**}, [\tau,t])$ vers $Z_{\theta_o}(\theta^{**}, [\tau,t])$ dans $\mathcal{C}(Kx\,[\tau,\tau+\delta])$ implique l'existence de n_1 tel que : $\forall n \ge n_1$

$$P\left\{\sup_{\theta^{**}\in K}\ \sup_{\tau\le t\le \tau+\delta} |Z_{n,\theta_n^{**}}(\theta^{**};[\tau,t]) -1| > \varepsilon\right\} \le \varepsilon$$

L'existence de n_2 satisfaisant : $\frac{\bar{T}}{\nu_{n_2}} < \delta(\varepsilon)$ achève la démonstration du lemme.

Pour terminer la preuve de la convergence uniforme sur les compacts de la suite π_n, il suffit de montrer la convergence du second terme qui est contenu dans le

LEMME (2.7) :

Sous les mêmes hypothèses, on a la convergence étroite uniforme sur les compacts de :

i) $Y_n(T) = q \text{ Log } Z_{n,\theta_n^{**}}(-\sqrt{n}\, d_n ; [\tau, \tau + \frac{T}{\nu_n}])$ dans $\mathcal{C}(\mathbb{R}^+)$ vers

$Y(T) = W(T) - \frac{T}{2}$

ii) $R_n(\theta^*,T) = \dfrac{Z_{n,\theta_n^{**}}(\theta^* - \sqrt{n}\, d_n ; [\tau, \tau + \frac{T}{\nu_n}])}{Z_{n,\theta_n^{**}}(-\sqrt{n}\, d_n ; [\tau, \tau + \frac{T}{\nu_n}])}$ dans $\mathcal{C}(\mathbb{R}^p \times \mathbb{R}^+)$ vers 1.

Démonstration du lemme (2.7) :

Par définition, on a :

$$Y_n(T) = \int_\tau^{\tau + \frac{T}{\nu_n}} \text{Log } \frac{\alpha(s,\theta_n^{**} - d_n)}{\alpha(s,\theta_n^{**})}\, dN_n(s) - \int_\tau^{\tau + \frac{T}{\nu_n}} n[\alpha(s,\theta_n^{**} - d_n) - \alpha(s,\theta_n^{**})]ds$$

$$\text{et Log } R_n(\theta^*,T) = \int_\tau^{\tau + \frac{T}{\nu_n}} \text{Log } \frac{\alpha(s,\theta_n^{**} - d_n + \frac{\theta^*}{\sqrt{n}})}{\alpha(s,\theta_n^{**} - d_n)}\, dN_n(s) - \int_\tau^{\tau + \frac{T}{\nu_n}} n[\alpha(s,\theta_n^{**} - d_n + \frac{\theta^*}{\sqrt{n}}) - \alpha(s,\theta_n^{**} - d_n)]\, ds$$

Un développement de Taylor au 2ème ordre au voisinage de $\theta_n^{**} - d_n$ pour Log R_n et l'expression des deux premiers moments du processus de Poisson N_n suffisent à montrer que pour un compact $K \times [0, \overline{T}]$ de $\mathbb{R}^{p+1}$

$$E_{\theta_n^{**}} \sup_{\theta^{**} \in K} \sup_{0 \leq t \leq \overline{T}} [\text{Log } R_n(\theta^*,T)]^2 \xrightarrow[n \to \infty]{} 0$$

Le processus $Y_n(.)$ est à accroissements indépendants et sa convergence vers $Y(.)$ dans $\mathcal{C}(\mathbb{R}^+)$ résulte de la convergence vers : $-\frac{T}{2}$ de son espérance:

$$E\, Y_n(T) = n \int_\tau^{\tau + \frac{T}{\nu_n}} [\text{Log } \frac{\alpha(s,\theta_n^{**} - d_n)}{\alpha(s,\theta_n^{**})} - \{\frac{\alpha(s,\theta_n^{**} - d_n)}{\alpha(s,\theta_n^{**} - d_n)} - 1\}]\ \alpha(s,\theta_n^{**})\, ds$$

et de la convergence en même temps que la majoration uniforme en n des variances, ce qui assure la tension :

$$V(Y_n(T) - Y_n(T')) = n \int_{\tau + \frac{T'}{\nu_n}}^{\tau + \frac{T}{\nu_n}} [\text{Log } \frac{\alpha(s,\theta_n^{**} - d_n)}{\alpha(s,\theta_n^{**})}]^2\ \alpha(s,\theta_n^{**})\, ds$$

Pour δ arbitrairement petit, il existe n_o tel que : $\forall n \geq n_o$,

$$V\,(Y_n(T)-Y_n(T')) \leq n.\,\frac{T-T'}{\nu_n}\,.\quad \sup_{\tau\leq s\leq\tau+\delta}\ \sup_{\|\theta\|\leq\delta}\left[{}^t d_n\ \frac{\partial}{\partial\theta}\ 2\sqrt{\alpha(s,\theta_o+\theta)}\right]^2$$

$$\leq B(T-T')$$

en utilisant la condition F.

Pour achever la démonstration du théorème (2.6), il nous faut montrer la convergence dans $\mathscr{C}_o(\mathbb{R}^{2p+1})$ de la suite π_n, c'est-à-dire le lemme suivant :

LEMME (2.8) :

Sous les conditions du théorème (2.6), on a :

$\forall\varepsilon > 0$, $\exists\overline{T} > 0$ tel que :

$$P\ \{\ \sup_{\theta^*\ \mathbb{R}^p}\ \sup_{\theta^{**}\ \mathbb{R}^p}\ \sup_{T\geq\overline{T}}\ \pi_n(\theta^*,\theta^{**},T) \geq \varepsilon\} \leq \varepsilon \qquad (*)$$

$\forall\varepsilon > 0$, $\forall\overline{T} > 0$, $\exists M$ tel que

$$P\{\ \sup_{\|\theta^*\|\geq M}\ \sup_{\|\theta^{**}\|\geq M}\ \sup_{0\leq T\leq\overline{T}}\ \pi_n(\theta^*,\theta^{**},T) \geq \varepsilon\} \leq \varepsilon \qquad (**)$$

Démonstration :

La décomposition de π_n en 3 processus indépendants :

$$\pi_n(\theta^*,\theta^{**},T)^{\frac{1}{q}} = Z_{n,\theta_n^*}(\theta^*,[0,\tau]).Z_{n,\theta_n^{**}}(\theta^*-\sqrt{n}\,d_n;[\tau,\tau+\frac{T}{\nu_n}]).Z_{n,\theta_n^{**}}(\theta^{**};[\tau+\tau\frac{T}{\nu_n},1])$$

et le découpage de la région correspondant à l'inégalité (**) en deux parties :

$$B_1 = \{(\theta^*,\theta^{**},T),\ \|\theta^*-\sqrt{n}\,d_n\| \leq 2\|\sqrt{n}\,d_n\|\ ,\ \|\theta^*\|\geq M,\ \|\theta^{**}\|\geq M,\ 0\leq T\leq\overline{T}\}$$

$$B_2 = \{(\theta^*,\theta^{**},T),\ \|\theta^*-\sqrt{n}\,d_n\| \geq 2\|\sqrt{n}\,d_n\|\ ,\ \|\theta^{**}\|\geq M,\ 0\leq T\leq\overline{T}\}$$

montrent qu'il suffit pour prouver (**) de vérifier les majorations suivantes :

$$P\ \{\sup_{B_1}\ Z_{n,\theta_n^{**}}\ (\theta^* - \sqrt{n}\ d_n,\ [\tau, \tau + \frac{T}{\nu_n}]) > A(\varepsilon)\} \leq \frac{\varepsilon}{5} \qquad (**.1)$$

$$P\{\sup_{\|\theta^*\| \geq M}\ Z_{n,\theta_n^*}\ (\theta^*, [0,\tau]) > \frac{\varepsilon}{A(\varepsilon)}\} \leq \frac{\varepsilon}{5} \qquad (**.2)$$

$$P\{\sup_{\|\theta^{**}\| > M}\ \sup_{0 \leq T \leq \bar{T}}\ Z_{n,\theta_n^{**}}\ (\theta^{**};\ [\tau + \frac{T}{\nu_n},\ 1]) > 1\} \leq \frac{\varepsilon}{5} \qquad (**.3)$$

$$P\ \{\sup_{B_2}\ Z_{n,\theta_n^{**}}\ (\theta^* - \sqrt{n}\ d_n;\ [\tau, \tau + \frac{T}{\nu_n}]) > A_n(\varepsilon)\} \leq \frac{\varepsilon}{5} \qquad (**.4)$$

$$P\ \{\sup_{B_2}\ Z_{n,\theta_n^*}\ (\theta^*, [0,\tau]) > \frac{\varepsilon}{A_n(\varepsilon)}\} \leq \frac{\varepsilon}{5} \qquad (**.5)$$

La première inégalité est satisfaite pour $A(\varepsilon)$ assez grand, quel que soit M à cause de la convergence sur les compacts de la suite de processus $Z_{n,\theta_n^{**}}\ (. - \sqrt{n}\ d_n,\ [\tau,\ \tau + \frac{\cdot}{\nu_n}])$ qui se montre comme le lemme (2.7).

La convergence de la suite $Z_{n,\theta_n^*}\ (.\ ;\ [0,\tau]\ )$ dans $\mathscr{C}_0(\mathbb{R}^p)$ et de la suite $Z_{n,\theta_n^{**}}(.,\ [.,1])$ dans $\mathscr{C}_0(\mathbb{R}^p \times [\tau\ ,\ \frac{\tau + 1}{2}])$ établie par les propositions (2.2) ou (2.3) implique l'existence pour $A(\varepsilon)$ fixé de M réalisant les inégalités (**.2) et (**.3).

L'inégalité (**.4) repose sur la convergence vers 0 de la suite $\frac{1}{\text{Log}\, A_n} \times \text{Log}\ Z_{n,\theta_n^{**}}\ (. - \sqrt{n}\ d_n, [\tau\ ,\ \tau + \frac{T}{\nu_n}])$ en dehors d'un compact fixe en θ : celle-ci se montrant exactement comme la proposition (2.3) pour ce processus nouvellement normalisé : $A_n = (\sqrt{n}\ d_n)^A$ avec A assez grand suffit

Enfin l'inégalité (**.5) s'apuie sur la vitesse de convergence vers 0 de la suite de processus $Z_{n,\theta_n^*}\ (.\ ;\ [0,\tau])$ en dehors d'un compact qui tend vers l'infini : $\|\theta^*\| > \sqrt{n}\ d_n$, d'après le lemme (2.5), il suffit de prendre

$$A_n = o\left((\sqrt{n}\ d_n)^{\frac{\nu' - \frac{p}{2}}{q'}} \right)\ ;$$

pour ν' assez grand, les inégalités (**.4) et **.5) sont satisfaites simultanément.

Pour prouver l'inégalité (*) qui donne la majoration du processus π_n en dehors d'un compact en T, nous partageons encore la région où nous prenons le suprémum en plusieurs parties où les 3 processus indépendants constituant π_n sont plus faciles à contrôler :

$$B_1' = \{(\theta^*,\theta^{**},T),\ \overline{T} \leq T \leq \frac{\nu_n(1-\tau)}{2}\ ,\ \|\theta^*\| \leq 2\ \|\sqrt{n}\ d_n\|\ \}$$

$$B_2' = \{(\theta^*,\theta^{**},T),\ \overline{T} \leq T \leq \frac{\nu_n(1-\tau)}{2}\ ,\ \|\theta^*\| \geq 2\ \|\sqrt{n}\ d_n\|\ \}$$

$$B_3' = \{(\theta^*,\theta^{**},T)\ ,\ \frac{\nu_n(1-\tau)}{2} \leq T \leq \nu_n(1-\tau)\ ,\ \|\theta^*\| \leq \frac{1}{2}\|\sqrt{n}\ d_n\|\ \}$$

$$B_4' = \{(\theta^*,\theta^{**},T),\ \frac{\nu_n(1-\tau)}{2} \leq T < \nu_n(1-\tau)\ ,\ \|\theta^*\| \geq \frac{1}{2}\|\sqrt{n}\ d_n\|\ \}$$

Sur B_1', la suite de processus $Z_{n,\theta_n^*}(.\ ;[0,\tau])$ est bornée en probabilité ainsi que la suite $Z_{n,\theta_n^{**}}(.;[\tau + \frac{\cdot}{\nu_n}\ ,\ 1])$ grâce à la convergence dans $\mathcal{C}_0$ de ces processus ; la convergence vers 0 de la suite $Z_{n,\theta_n^*}(.-\sqrt{n}\ d_n;\ [\tau,\ \tau+\frac{\cdot}{\nu_n}])$ sur un compact fixe en θ et en dehors d'un compact en T tendant vers l'infini assure le contrôle de π_n sur B_1' pour $\overline{T}$ assez grand.

Sur B_2', on a encore la suite $Z_{n,\theta_n^{**}}(.\ ;[\tau + \frac{\cdot}{\nu_n},\ 1])$ bornée en probabilité mais les deux autres suites $Z_{n,\theta_n^*}(.\ ;\ [0,\tau])$ et $Z_{n,\theta_n^{**}}(.-\sqrt{n}\ d_n,[\tau,\tau+\frac{\cdot}{\nu_n}])$ tendent vers zéro en dehors de leurs compacts respectifs pourvu que $\overline{T}$ soit choisi suffisamment grand.

Sur B_3', la suite $Z_{n,\theta_n^*}(.\ ;\ [0,\tau])$ est bornée en probabilité ; la suite $Z_{n,\theta_n^{**}}(.-\sqrt{n}\ d_n\ ;\ [\tau,\ \tau+\frac{\cdot}{\nu_n}])$ est majorée en probabilité par $\frac{1}{A_n}$ grâce à la proposition (2.3) et la suite $Z_{n,\theta_n^{**}}(.\ ;\ [\tau+\frac{\cdot}{\nu_n},1])$ est majorée en probabilité par $\frac{\varepsilon}{A(\varepsilon)}\ .\ A_n$ grâce au lemme (2.5).

Sur B_4', la suite $Z_{n,\theta_n^*}(.,\ [0,\tau])$ est majorée en probabilité par $\frac{1}{A_n}$, la suite $Z_{n,\theta_n^{**}}(.\ -\sqrt{n}\ d_n,\ [\tau,\tau+\frac{\cdot}{\nu_n}])$ est bornée en probabilité et $Z_{n,\theta_n^{**}}(.;[\tau+\frac{\cdot}{\nu_n},1])$ est encore majoré en probabilité par $\frac{\varepsilon}{A(\varepsilon)}\ .\ A_n$.

III. RUPTURE DANS DES MODELES DE PROCESSUS PONCTUELS PLUS GENERAUX.

Jusqu'à maintenant, nous avons envisagé des modèles de processus de Poisson sur la droite réelle d'intensité paramétrique avec un changement de paramètre éventuel au cours du temps. Les méthodes que nous avons développées pour détecter une rupture et pour en estimer les paramètres sont celles du maximum de vraisemblance, elles devraient donc s'étendre à beaucoup de situations où il est possible d'écrire des rapports de vraisemblance et de les approximer.

3.1. Processus de Poisson spatial homogène.

La structure temporelle du processus de Poisson homogène sur la droite étudié au paragraphe I facilitait la recherche du changement de paramètre, l'alternative à l'homogénéité sur tout l'intervalle d'observation étant l'existence d'un ou plusieurs points partageant la durée d'échantillonnage en intervalles où le processus est homogène.

Pour un processus de Poisson spatial, les ruptures seront caractérisées par des frontières séparant l'espace d'échantillonnage (par exemple : un pavé) en plusieurs régions où le processus sera homogène : l'estimation des frontières par le maximum de vraisemblance ne posera pas de problèmes à partir du moment où on connaît une forme paramétrique de ces frontières inconnues. Par exemple, on peut chercher des droites parallèles à une droite donné (ou hyperplans), des droites quelconques, des cercles concentriques ou non (ou sphères), ... partitionnant le pavé d'observation (ou la boule). Les lois asymptotiques s'obtiendront de façon identique au cas unidimensionnel.

3.2. Processus de Poisson composé.

Reprenons le modèle étudié au paragraphe II : processus de Poisson sur la droite d'intensité paramétrique ; si à chaque instant d'observation est associé une variable aléatoire (ou marque) on définit ainsi un processus de Poisson composé. Lorsque ces variables aléatoires sont indépendantes, identiquement distribuées, et indépendantes du processus de comptage, l'étude des modèles avec rupture sur le paramètre de l'intensité ou sur la loi commune des marques découle directement des résultats précédents et de [15] . La différence entre le processus de Poisson composé et le modèle consistant en une suite de variables indépendantes, identiquement distribuées, indicée chronologiquement réside dans le fait que le premier modèle contient en plus un échantillonnage du temps.

Ce point de vue devient intéressant pour l'étude des modèles de régression: si on observe une suite de variables aléatoires Y_i que l'on régresse sur des vecteurs aléatoires X_i où les couples (X_i,Y_i), $i=1,\ldots,n$ sont indépendants ; les variables X_i peuvent être rangées par ordre croissant de l'une des coordonnées et on peut envisager des changements de paramètre suivant la valeur de cette coordonnée. Les points (X_i,Y_i) observés peuvent être considérés comme l'échantillon d'un processus de Poisson composé dont les marques seraient les autres coordonnées de X et Y. La convergence des processus de vraisemblance dans ce cadre est immédiate et on en déduit la loi limite des estimateurs du maximum de vraisemblance comme aux paragraphes I et II, que l'amplitude de la rupture tende asymptotiquement vers 0 ou non. Cela complète l'étude faite dans [15] où les régresseurs étaient réparties de façon contrôlée sur la droite : en fait, équirépartis sur $[0,1]$. Un modèle généralisant le processus de Poisson composé où des ruptures au cours du temps sont envisageables est le :

3.3. Processus de saut markovien :

Les instants de sauts et les marques ont une loi dépendant de la marque précédente : de façon plus précise si $T_1,\ldots,T_n,\ldots$ désignent les instants successifs de saut et $Y_1,\ldots,Y_n,\ldots$ les marques associées, la loi du processus est définie par le fait que conditionnellement à $(T_1,\ldots,T_n, Y_1,\ldots,Y_n)$, les variables Y_{n+1} et $T_{n+1}-T_n$ sont indépendantes et de loi respective : $p(.,Y_n)$ et exponentielle de paramètre $\lambda(Y_n)$. Les processus de vie et de mort où les Y_i sont à valeurs dans $\mathbb{N}$ avec : $p(y_n+1, y_n) + p(y_n-1, y_n) = 1$, les processus de queues M/M avec un ou plusieurs serveurs en sont des cas particuliers importants.

Pour ces modèles, l'écriture des vraisemblances est immédiate dès que les lois sont paramétriques : $p(.,y_n,\theta_1)$ et $\lambda(y_n,\theta_2)$, et les lois asymptotiques des estimateurs du maximum de vraisemblance lorsqu'il y a changement de l'un des paramètres θ_1 ou θ_2 ou des deux se calculent encore.

En fait, l'intensité peut prendre des formes plus générales : $\Lambda(t,\theta)$ où Λ est un processus $\mathcal{F}_t$-prévisible positif, $\mathcal{F}_t$ étant une famille de tribus définissant l'histoire jusqu'au temps t, éventuellement plus grosse que l'histoire interne du processus ponctuel N notée $\mathcal{F}^N_t$. A partir du moment où on considère des mesures absolument continues, les rapports de vraisemblance s'écrivent par un théorème analogue au théorème de Girsanov (voir [9]). Une étude asymptotique des estimateurs du maximum de vraisemblance de θ a été faite par OGATA [33] pour des intensités Λ très régulières. Des changements de paramètre peuvent être envisagés par exemple pour des intensités qui se factorisent

$$\Lambda(t,\theta) = \alpha(t,\theta)\, Y(t)$$

où α est une fonction déterministe (paramétrique) et Y un processus $\mathcal{F}_t$-prévisible et pour des processus "self exciting" (voir [19], [35]) où l'intensité se met sous la forme :

$$\Lambda(t,\theta) = \Lambda(0,\theta) + \int_0^t \alpha(s,\theta)\, dN(s).$$

Lorsque le modèle n'est plus paramétrique, il faut évidemment envisager d'autres méthodes que celles de vraisemblance pour détecter et estimer des changements de modèles.

3.4. Rupture dans des modèles de type non paramétrique.

Une première situation de ce type est l'observation d'un processus ponctuel stationnaire du 2ème ordre. Une rupture dans les caractéristiques du 2ème ordre se traduira par un changement de mesure spectrale. Une étude analogue à celle conduite par D. PICARD [36] en temps discret, pourrait être menée à bien ; toujours en comparant les fonctions de répartition spectrales estimées avant et après rupture, on peut tester la stationnarité du processus.

Une seconde situation de type non paramétrique est celle où l'intensité est proportionnelle à un processus prévisible Y :

$$\Lambda(t) = \alpha(t) \,.\, Y(t)$$

avec α complètement inconnue (voir par exemple AALEN [1] , [2] , ANDERSEN [3] , ANDERSEN-GILL [4]), une rupture dans un tel modèle pourrait être un saut de la fonction α supposée régulière ailleurs. Nous sommes alors confrontés à des problèmes analogues à ceux de n-échantillon avec saut de densité.

Enfin, nous pourrions étudier des ruptures dans des modèles de survie (voir COX [12],[13], OAKES [32] , KALBFLEISH-PRENTICE [25]) où le taux de hasard du i^e individu est décrit par :

$$\Lambda_i(t) = \alpha(t)\, e^{t_\beta Z_i(t)}$$

avec Z_i dépendant ou non du temps jouant le rôle de régresseur et $\alpha(.)$ paramétrique ou non étant considéré comme fonction fantôme. L'histoire jusqu'au temps t étant résumée par l'ensemble des "individus à risque" au temps t : $\mathcal{R}(t)$.

L'inférence sur β est très souvent faite à partir de la vraisemblance partielle ([13]) :

$$\frac{e^{t_\beta Z_{i_j}(t_j)}}{\sum_{i \in \mathcal{R}(t_j)} e^{t_\beta Z_i(t_j)}}$$

si i_j est le numéro de l'individu qui manque au temps t_j ; une étude asymptotique des méthodes de vraisemblance partielle avec rupture sur le paramètre β pourrait être faite comme elle l'a été en l'absence de rupture par TSIATIS [42] . Une étude globale en tenant compte la fonction fantôme α peut aussi être envisagée : dans [16], FRIEDMAN fait un découpage fixe de l'intervalle d'observation en sous-intervalles où α est supposée constante : une alternative serait de supposer α constante (ou paramétrique) par morceaux mais d'estimer les instants où il y a changement de paramètres.

BIBLIOGRAPHIE

[1] AALEN O. : " Nonparametric inference for a family of counting processes ". The Annals of Statistics (1978) vol. 6, p. 701-726.

[2] AALEN O. : " A model for nonparametric regression analysis of counting processes ". Springer Lect. Notes in Statistics, n°2, p. 1-25 (WISLA 1978).

[3] ANDERSEN P.K. : " Testing goodness of fit of Cox's regression and life model ". Biometrics (1982) vol. 38, p. 67-77.

[4] ANDERSEN P.K. - GILL R. : " Cox's regression model for counting processes : a large sample study ". Research Report 81/6. Statistical Research Unit, Danish Medical and Social Science Research Councils.

[5] BASAWA I.V. : " Remarks on Bahadur optimality of conditional tests ". The Annals of Statistics (1980) vol. 8, p. 1382-1387.

[6] BASAWA I.V. - PRAKASA RAO B.L.S. : " Statistical inference for stochastic processes ". Academic Press (1980).

[7] BERK R.H. - JONES D.H. : " Goodness of fit test statistics that dominate the Kolmogorov statistics ". Z. WAhrsch. verw. Gebiete (1979) vol.47, p. 47-59.

[8] BIRNBAUM Z.W. - MARSHALL A.W. : " Some multivariate Chebyschev inequalities with extensions to continuous parameter processes ". A.M.S. (1961) vol. 32, p. 687-703.

[9] BREMAUD P. - JACOD J. : " Processus ponctuels et martingales : résultats récents sur la modélisation et le filtrage ". Adv. Appl. Prob. (1977) vol. 9, p. 362-416.

[10] BROWN L.D. : " Non-local asymptotic optimality of appropriate likelihood ratio tests ". A.M.S. (1971) vol. 42, p. 1206-1240.

[11] BROWN M. : " Statistical analysis of non-homogeneous Poisson processes ". Dans Stochastic Point Processes (éd. LEWIS) (1972) Wiley, p. 67-89.

[12] COX D.R. : " Regression models and life-tables (with discussion) ". J.R.S.S. B (1972) vol. 34, p. 187-220.

[13] COX D.R. : " Partial likelihood ". Biometrika (1975) vol. 62, p. 269-276.

[14] DESHAYES J. - PICARD D. : " Principe d'invariance sur les processus de vraisemblance ". Prépublication d'Orsay (1982) t. 10.

[15] DESHAYES J. - PICARD D. : " Rupture dans les modèles de régression : lois asymptotiques des tests et estimateurs du maximum de vraisemblance ". Prépublication d'Orsay (1982) t. 20.

[16] FRIEDMAN M. : " Piecewise exponential models for survival data with covariates ". The Annals of Statistics (1982) vol. 10, p. 101-113.

[17] GOVINDARAJULU Z. - LE CAM L. - RACHAVACHARI M. : " Generelizations of theorems of Chernoff and Savage on the asymptotic normality of tests statistics ". Vth Berkeley Symposium I. (1967) p. 609-638.

[18] GRANDELL J. : " Doubly stochastic Poisson processes ". Lecture Notes in Mathematics n° 529.

[19] HAWKES A.G. : " Spectra of some self-exciting and mutually exciting point processes ". Biometrika (1971) vol. 58, p. 83-90.

[20] HINKLEY D.V. : " Inference about the change-point in a sequence of random variables ". Biometrika (1970) vol. 57, p. 1-17.

[21] HINKLEY D.V. - HINKLEY E.A. : " Inference about the change-point in a sequence of binomial variables ". Biometrika (1970) vol. 57, p. 477-488.

[22] IBRAGIMOV I.A. - HAS'MINSKII R.Z. : " Local asymptotic normality for non identically distributed observations ". T.P.A. (1975) vol. 20, p. 246-260.

[23] IBRAGIMOV I.A. - HAS'MINSKII R.Z. : " Statistical estimation - asymptotic theory ". Springer Verlag, Applications of Mathematics n° 16.

[24] JOHANSEN S. : " An extension of Cox's regression model ". Preprint 81/11 Institute of Mathematical Statistics. University of Copenhagen.

[25] KALBFLEISH J.D. - PRENTICE R.L. : " The statistical analysis of Failure time date ". Wiley (1980).

[26] KEIDING N. : " Maximum likelihood estimation in the birth-and-death process ". The Annals of Statistics (1975) vol. 3, p. 363-372.

[27] KUTOYANTS Y.A. : " Multidimensional parameter estimation of the intensity function of inhomogeneous Poisson processes ". Problems of Control and Information Theory (1982) vol. 11, p. 325-334.

[28] LEWIS P.A. : " Recent results in the statistical analysis of univariate point processes ". Dans Stochastic Point Porcesses (ed. LEWIS) (1972) Wiley, p. 1-54.

[29] MATTHEWS D.E. - FAREWELL V.T. : " On testing for a constant hazard against a change-point alternative ". Biometrics (1982) vol. 38, p. 463-468.

[30] MILNE R.K. - WESTCOTT M. : " Further results for Gauss-Poisson processes ". Adv. Appl. Prob. (1972) vol. 4, p. 151-176.

[31] NOVIKOV A.A. : " On discontinuous martingales ". T.P.A. (1975) vol. 20, p. 11-26.

[32] OAKES D. : " Survival times : aspects of partial likelihood ". Inst. Stat. Review (1981) vol. 49, p. 235-264.

[33] OGATA Y. : " The asymptotic behaviour of maximum likelihood estimators for stationary point processes ". Ann. Inst. Stat. Math. (1978) vol. 30 part A, p. 243-261.

[34] OGATA Y. - AKAIKE H. : " On linear intensity models for mixed doubly stochastic Poisson and self exciting point processes ". J.R.S.S. B (1982) vol. 44, p. 102-107.

[35] OGATA Y. - AKAIKE H. - KATSURA K. : " The application of linear intensity models to the investigation of causal relations between a point process and another stochastic process ". Ann. Inst. Stat. Math. (1982) vol. 34 B p. 373-387.

[36] PICARD D. : " Rupture dans une série chronologique (tests et estimation) ". Prépublication d'Orsay (1982).

[37] PRENTICE R.L. : " Covariate measurement errors and parameter estimation in a failure time regression model ". Biometrika (1982) vol. 69, p. 331-342.

[38] PYKE R. - SHORACK G. : " Weak convergence of a two-sample empirical process and a new approach to Chernoff-Savage theorems ". A.M.S. (1968) vol. 39, p. 755-771.

[39] REBOLLEDO R. : " Central limit theorems for local martingales ". Z. Wahrscheinlichkeitstheorie verw. Gebiete (1980) vol. 51, p. 269-286.

[40] SNYDER D.L. : " Random point processes ". Wiley (1975).

[41] STONE M. : " Large deviations of empirical probability measures ". The Annals of Statistics (1974) vol. 2, p. 362-366.

[42] TSIATIS A.A. : " A large sample study of Cox's regression model ". The Annals of Statistics (1981) vol. 9, p. 93-108.

[43] VERE-JONES D. - OZAKI T. : " Estimation applied to earthquake data ". Ann. Inst. Stat. Math. (1982) vol. 34, p. 189-207.

N° D'IMPRESSION 600
2ème TRIMESTRE 1983

RESUME.

Cette thèse présente l'étude asymptotique approfondie de procédures de détection et d'estimation de ruptures dans des modèles statistiques (indicés chronologiquement). Pour des cadres statistiques variés (modèles paramétriques ou non, observations indépendantes ou non, temps discret ou continu), on trouvera une présentation globale des méthodes classiques, la description de nouvelles procédures, ainsi qu'une comparaison systématique de celles-ci au moyen de simulations et d'une étude asymptotique. L'optimalité quasi-générale des méthodes de vraisemblance nous a amené à en approfondir l'étude : un théorème de convergence du processus de vraisemblance à double indice (paramètre, instant de rupture) vers un processus s'exprimant à l'aide d'un brownien avec dérive, fournit une étude globalisante et nouvelle : en découlent en particulier la liaison entre les paramètres instant et amplitude de la rupture, la loi limite des statistiques de test de vraisemblance renormalisées et celle des estimateurs du maximum de vraisemblance.

MOTS CLEFS : Rupture de modèles, régression, mouvement brownien, Poisson (processus de), processus empirique, test non paramétrique. Vraisemblance : processus, test, estimateurs, lois asymptotiques, grandes déviations.